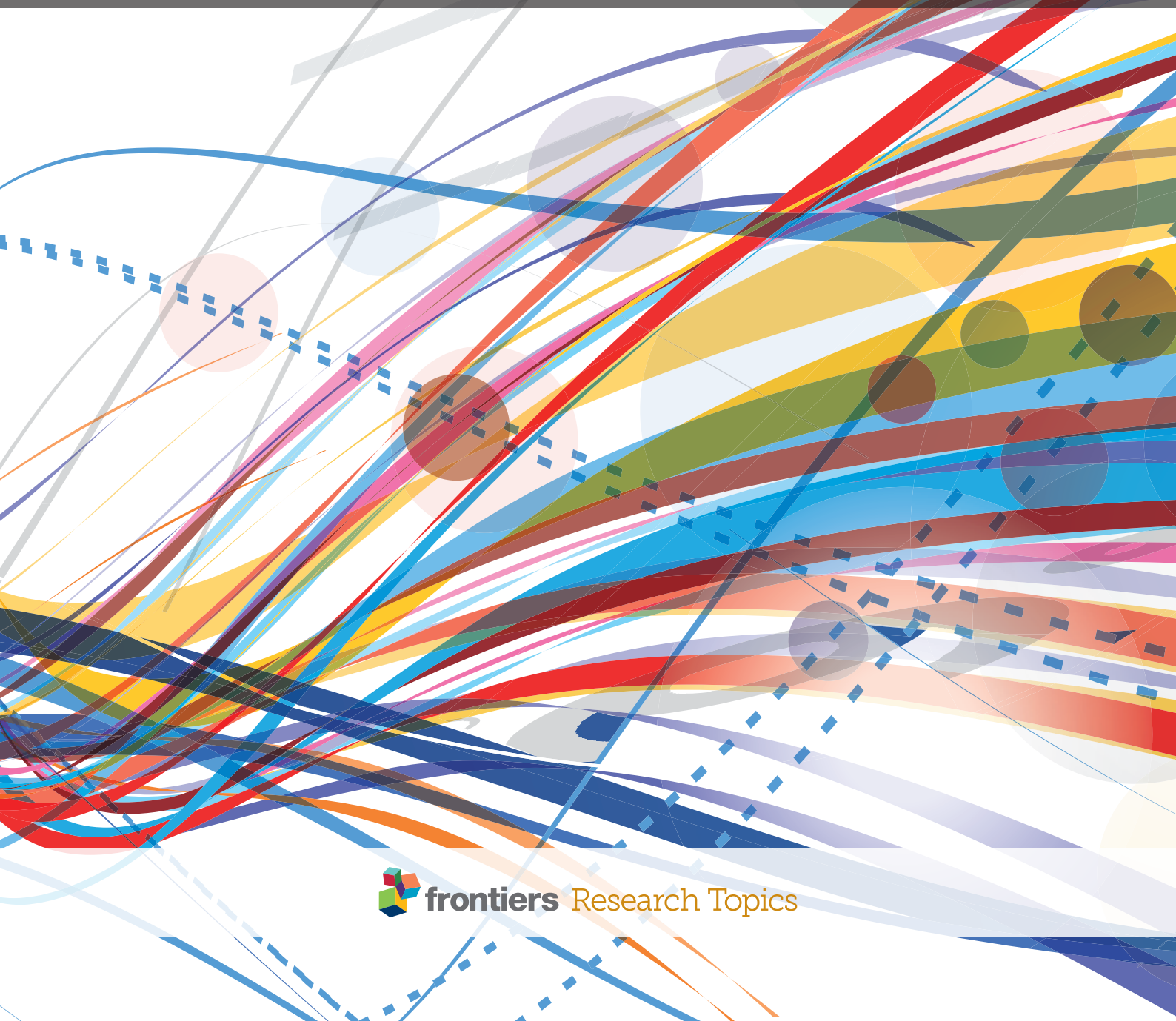
DEVELOPMENT OF EXECUTIVE FUNCTION DURING CHILDHOOD

EDITED BY: Yusuke Moriguchi, Philip D. Zelazo and Nicolas Chevalier
PUBLISHED IN: Frontiers in Psychology



frontiers Research Topics

frontiers

Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714
ISBN 978-2-88919-800-9
DOI 10.3389/978-2-88919-800-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

DEVELOPMENT OF EXECUTIVE FUNCTION DURING CHILDHOOD

Topic Editors:
Yusuke Moriguchi, Joetsu University of Education, Japan
Philip D. Zelazo, University of Minnesota, USA
Nicolas Chevalier, University of Edinburgh, UK

Executive function refers to the goal-oriented regulation of one's own thoughts, actions, and emotions. Its importance is attested by its contribution to the development of other cognitive skills (e.g., theory of mind), social abilities (e.g., peer interactions), and academic achievement (e.g., mathematics), and by the consequences of deficits in executive function (which are observed in wide range of developmental disorders, such as attention-deficit hyperactivity disorder and autism). Over the last decade, there have been growing interest in the development of executive function, and an expanding body of research has shown that executive function develops rapidly during the preschool years, with adult-level performance being achieved during adolescence or later.

This recent work, together with experimental research showing the effects of interventions targeting executive function, has yielded important insights into the neurocognitive processes underlying executive function. Given the complexity of the construct of executive function, however, and the multiplicity of underlying processes, there are often inconsistencies in the way that executive function is defined and studied. This inconsistency has hampered communication among researchers from various fields.

This Research Topic is intended to bridge this gap and provide an opportunity for researchers from different perspectives to discuss recent advances in understanding childhood executive function. Researchers using various methods, including, behavioral experiments, neuroimaging, eye-tracking, computer simulation, observational methods, and questionnaires, are encouraged to contribute original empirical research. In addition to original empirical articles, theoretical reviews and opinions/perspective articles on promising future directions are welcome. We hope that researchers from different areas, such as developmental psychology, educational psychology, experimental psychology, neuropsychology, neuroscience, psychiatry, computational science, etc., will be represented in the Research Topic.

Citation: Moriguchi, Y., Zelazo, P. D., Chevalier, N., eds. (2016). Development of Executive Function during Childhood. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-800-9

Table of Contents

06 ***Editorial: Development of Executive Function during Childhood***
Yusuke Moriguchi, Nicolas Chevalier and Philip D. Zelazo

08 ***Primed to be inflexible: the influence of set size on cognitive flexibility during childhood***
Lily FitzGibbon, Lucy Cragg and Daniel J. Carroll

21 ***"Trick or treat": the influence of incentives on developmental changes in feedback-based learning***
Kerstin Unger, Berit Greulich and Jutta Kray

30 ***Stuck in the moment: cognitive inflexibility in preschoolers following an extended time period***
Carolina Garcia and Anthony Steven Dick

36 ***An effect of inhibitory load in children while keeping working memory load constant***
Andy Wright and Adele Diamond

45 ***Age-related changes in the temporal dynamics of executive control: a study in 5- and 6-year-old children***
Joanna Lucenet and Agnès Blaye

56 ***Gaining control: changing relations between executive control and processing speed and their relevance for mathematics achievement over course of the preschool period***
Caron A. C. Clark, Jennifer Mize Nelson, John Garza, Tiffany D. Sheffield, Sandra A. Wiebe and Kimberly Andrews Espy

71 ***Longitudinal and concurrent links between memory span, anxiety symptoms, and subsequent executive functioning in young children***
Laura Visu-Petra, Oana Stanciu, Oana Benga, Mircea Miclea and Lavinia Cheie

92 ***Inhibitory processes in toddlers: a latent-variable approach***
Elena Gandolfi, Paola Viterbori, Laura Traverso and M. Carmen Usai

103 ***Psychometric properties and convergent and predictive validity of an executive function test battery for two-year-olds***
Hanna Mulder, Huub Hoofs, Josje Verhagen, Ineke van der Veen and Paul P. M. Leseman

120 ***Development of neural mechanisms of conflict and error processing during childhood: implications for self-regulation***
Purificación Checa, M. C. Castellanos, Alicia Abundis-Gutiérrez and M. Rosario Rueda

133 ***Stability of executive function and predictions to adaptive behavior from middle childhood to pre-adolescence***
Madeline B. Harms, Vivian Zayas, Andrew N. Meltzoff and Stephanie M. Carlson

144 ***Trail making test performance in youth varies as a function of anatomical coupling between the prefrontal cortex and distributed cortical regions***
Nancy Raitano Lee, Gregory L. Wallace, Armin Raznahan, Liv S. Clasen and Jay N. Giedd

158 ***The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective***
Gina Austin, Karoline Groppe and Birgit Elsner

169 ***Development of reference assignment in children: a direct comparison to the performance of cognitive shift***
Taro Murakami and Kazuhide Hashiya

176 ***The role of executive functions in the pragmatic skills of children age 4–5***
Bénédicte Blain-Brière, Caroline Bouchard and Nathalie Bigras

190 ***The union of narrative and executive function: different but complementary***
Margaret Friend and Raven Phoenix Bates

202 ***Sharing and giving across adolescence: an experimental study examining the development of prosocial behavior***
Berna Güroğlu, Wouter van den Bos and Eveline A. Crone

215 ***The early development of executive function and its relation to social interaction: a brief review***
Yusuke Moriguchi

221 ***Relations between executive function and emotionality in preschoolers: Exploring a transitive cognition–emotion linkage***
David E. Ferrier, Hideko H. Bassett and Susanne A. Denham

233 ***Child temperamental reactivity and self-regulation effects on attentional biases***
Georgiana Susa, Oana Benga, Irina Pitica and Mircea Miclea

244 ***Executive function and food approach behavior in middle childhood***
Karoline Groppe and Birgit Elsner

256 ***Getting the right grasp on executive function***
Claudia L. R. Gonzalez, Kelly J. Mills, Inge Genee, Fangfang Li, Noella Piquette, Nicole Rosen and Robbin Gibb

267 ***The role of executive function in children's source monitoring with varying retrieval strategies***
Becky Earhart and Kim P. Roberts

279 ***Measuring inhibitory control in children and adults: brain imaging and mental chronometry***
Olivier Houdé and Grégoire Borst

286 ***Executive functioning and reading achievement in school: a study of Brazilian children assessed by their teachers as "poor readers"***
Pascale M. J. Engel de Abreu, Neander Abreu, Carolina C. Nikaedo, Marina L. Puglisi, Carlos J. Tourinho, Mônica C. Miranda, Debora M. Befi-Lopes, Orlando F. A. Bueno and Romain Martin

300 ***Predictors of early growth in academic achievement: the head-toes-knees-shoulders task***
Megan M. McClelland, Claire E. Cameron, Robert Duncan, Ryan P. Bowles, Alan C. Acock, Alicia and Megan E. Pratt

314 ***Sorting Test, Tower Test, and BRIEF-SR do not predict school performance of healthy adolescents in preuniversity education***
Annemarie Boschloo, Lydia Krabbendam, Aukje Aben, Renate de Groot and Jelle Jolles

322 ***Executive control training from middle childhood to adolescence***
Julia Karbach and Kerstin Unger

336 ***Predictors of cognitive enhancement after training in preschoolers from diverse socioeconomic backgrounds***
M. Soledad Segretin, Sebastián J. Lipina, M. Julia Hermida, Tiffany D. Sheffield, Jennifer M. Nelson, Kimberly A. Espy and Jorge A. Colombo

357 ***The potential adverse effect of energy drinks on executive functions in early adolescence***
Tamara Van Batenburg-Eddes, Nikki C. Lee, Wouter D. Weeda, Lydia Krabbendam and Mariette Huizinga

366 ***Less-structured time in children's daily lives predicts self-directed executive functioning***
Jane E. Barker, Andrei D. Semenov, Laura Michaelson, Lindsay S. Provan, Hannah R. Snyder and Yuko Munakata

382 ***Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks***
Virginia C. Mueller Gathercole, Enlli M. Thomas, Ivan Kennedy, Cynog Prys, Nia Young, Nestor Viñas Guasch, Emily J. Roberts, Emma K. Hughes and Leah Jones

396 ***ADHD among young adults born at extremely low birth weight: the role of fluid intelligence in childhood***
Ayelet Lahat, Ryan J. Van Lieshout, Saroj Saigal, Michael H. Boyle and Louis A. Schmidt

403 ***The development of cognitive control in children with chromosome 22q11.2 deletion syndrome***
Heather M. Shapiro, Flora Tassone, Nimrah S. Choudhary and Tony J. Simon

417 ***Investigating executive functions in children with severe speech and movement disorders using structured tasks***
Kristine Stadskleiv, Stephen von Tetzchner, Beata Batorowicz, Hans van Balkom, Annika Dahlgren-Sandberg and Gregor Renner

431 ***Age-related trends of inhibitory control in Stroop-like big–small task in 3 to 12-year-old children and young adults***
Yoshifumi Ikeda, Hideyuki Okuzumi and Mitsuru Kokubun

437 ***Age differences in high frequency phasic heart rate variability and performance response to increased executive function load in three executive function tasks***
Dana L. Byrd, Erin T. Reuther, Joseph P. H. McNamara, Teri L. DeLucca and William K. Berg

450 ***Task-evoked pupillometry provides a window into the development of short-term memory capacity***
Elizabeth L. Johnson, Alison T. Miller Singley, Andrew D. Peckham, Sheri L. Johnson and Silvia A. Bunge

frontiers in Psychology

EDITORIAL
published: 20 January 2016
doi: 10.3389/fpsyg.2016.00006

Editorial: Development of Executive Function during Childhood

Yusuke Moriguchi[1]*, *Nicolas Chevalier*[2] and *Philip D. Zelazo*[3]

[1] Department of School Education, Joetsu University of Education, Joetsu, Japan, [2] Department of Psychology, The University of Edinburgh, Edinburgh, UK, [3] Institute of Child Development, University of Minnesota, Minneapolis, MN, USA

Keywords: executive function, cognitive flexibility, inhibitory control, working memory, brain development, socio-emotional development, prefrontal cortex (PFC), cognitive development

The Editorial on the Research Topic

Development of Executive Function during Childhood

Executive function (EF) is one of the most rapidly expanding research fields in the developmental and cognitive sciences. The aim of this Research Topic is to present a broad sample of recent advances in understanding the development of EF. The 38 articles in this collection provide a unique, state-of-the-art tour of current, burning issues regarding executive function development, from cutting-edge research on the underpinning basic cognitive processes to the most promising applications in educational and clinical settings.

COGNITIVE PROCESSES OF EF DURING CHILDHOOD

EF involves several complex cognitive processes, including working memory, inhibitory control, and cognitive flexibility. The present papers shed new light on how these processes develop and how they are interrelated. Specifically, they clarify the conditions that modulate EF demands (FitzGibbon et al.; Unger et al.), how their effect can persist in time (Garcia and Dick), the specific executive processes (e.g., inhibitory control) at play in a given task (Wright and Diamond) and the specific age windows during which critical changes in EF engagement occur (Lucenet and Blaye). Furthermore, they provide new evidence that EF may develop through progressive differentiation of executive processes from more basic cognitive processes (e.g., processing speed and short-term memory; Clark et al.; Visu-Petra et al.) and of different forms of EF (e.g., cognitive "cool" EF vs. affective "hot" EF) (Gandolfi et al.; Mulder et al.). They further identify the brain correlates of EF development using EEG/ERP or MRI (Checa et al.; Harms et al.; Unger et al.), revealing for instance that anatomical coupling between the left prefrontal cortex and other distributed brain regions predicts behavioral performance (Lee et al.).

THE CRITICAL ROLE OF EF IN SOCIAL, EMOTIONAL, AND COGNITIVE DEVELOPMENT

The present papers also reveal or clarify the association of EF to a host of social and emotional processes including, for instance, theory of mind (Austin et al.), referent assignment (Murakami and Hashiya), conversational pragmatics (Blain-Brière et al.), narrative skills (Friend and Bates), prosocial behaviors (Güroğlu et al.), social interactions (Moriguchi), sensation seeking

OPEN ACCESS

Edited and reviewed by:
Jessica S. Horst, University of Sussex, UK

***Correspondence:**
Yusuke Moriguchi
moriguchi@juen.ac.jp

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *15 December 2015*
Accepted: *04 January 2016*
Published: *20 January 2016*

Citation:
Moriguchi Y, Chevalier N and Zelazo PD (2016) Editorial: Development of Executive Function during Childhood. Front. Psychol. 7:6. doi: 10.3389/fpsyg.2016.00006

(Harms et al.), emotional experience (Ferrier et al.), fear (Susa et al.), and emotional overeating (Groppe and Elsner). Impressively, these associations are often found over and above associations with IQ. Other findings highlight links between EF and motor function (Gonzalez et al.), source monitoring (Earhart and Roberts), and conceptual development (Houdé and Borst). These impressive findings highlight the foundational role that EF plays in goal-directed behavior across a wide range of domains and situations, and they underscore that the healthy development of EF skills is critical for both social-emotional and cognitive development. Indeed, they suggest that understanding the development of EF is absolutely key to understanding child development overall.

EF AND ACADEMIC ACHIEVEMENT

One of the most important foci in research on EF is the relation of EF development to school readiness and academic achievement. The studies included in this Research Topic provide further evidence of the predictive value of EF in academic learning, and in particular reading (Engel de Abreu et al.). Critically, they also clarify the discriminative importance of EF processes for children's mathematical learning, showing how the role of EF may increase from preschool to kindergarten (Clark et al.; McClelland et al.) and then wane in adolescence (Boschloo et al.). Such findings charting out the influence of EF on academic learning are essential to designing effective interventions that target strategic time points in development. Indeed, extant evidence suggests that such training programs can effectively enhance academic achievement (Karbach and Unger; Segretin et al.), although socio-environmental factors, such as housing conditions, may moderate the effects of cognitive interventions in children (Segretin et al.).

EXPERIENCES AFFECTING EF

Given the importance of EF for child development and academic achievement, several studies examined experiential influences that may affect its development. The findings suggest that some activities, such as regular energy drink consumption during adolescence, may *impair* EF (Batenburg-Eddes et al.), whereas others, such as time spent in non-structured activities, may *promote* it during childhood (Barker et al.). Meanwhile, the influence of other factors that have long been assumed to affect EF, in particular bilingualism, may have been overestimated in the past (Gathercole et al.). All these thought-provoking findings have important implications on societal choices and for policy makers.

EF DISORDERS

Just as EF appears to play an essential role in typical development, difficulties in EF are central features of several developmental disorders. The studies in this Research Topic contribute to clarifying the role of EF in ADHD symptoms (Lahat et al.), Chromosome 22q11.2 Deletion Syndrome (Shapiro et al.), and severe speech and motor impairments (Stadskleiv et al.).

MEASURING EF IN CHILDREN

Finally, advances in research on EF development rely critically on designing effective, valid, and reliable instruments and methodologies. The present papers greatly contribute to this effort by developing new EF tasks (Ikeda et al.) and showing how physiological measures, such as pupil dilation and phasic heart rate variability (HRV), can bring further insight on children's EF (Byrd et al.; Johnson et al.).

SUMMARY

The articles in this Research Topic demonstrate how considerations of both basic cognitive/biological processes and applied/clinical settings help to unify and extend our understanding of EF during childhood. They illustrate the large range of questions and debates that animate this particularly dynamic field. We hope that this Research Topic will be helpful to both novices and experts of EF development by providing an overview of the field and highlighting the most recent advances.

AUTHOR CONTRIBUTIONS

All authors drafted the manuscript, and provided critical revisions. All authors approved the final version of the manuscript for submission.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Moriguchi, Chevalier and Zelazo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 12 February 2014
doi: 10.3389/fpsyg.2014.00101

Primed to be inflexible: the influence of set size on cognitive flexibility during childhood

Lily FitzGibbon[1]*, Lucy Cragg[2] and Daniel J. Carroll[1]

[1] *Department of Psychology, University of Sheffield, Sheffield, UK*
[2] *School of Psychology, University of Nottingham, Nottingham, UK*

Edited by:
Nicolas Chevalier, University of Colorado Boulder, USA

Reviewed by:
Jutta Kray, Saarland University, Germany
Vanessa R. Simmering, University of Wisconsin Madison, USA

****Correspondence:***
Lily FitzGibbon, Department of Psychology, University of Sheffield, Western Bank, Sheffield S10 2TP, UK
e-mail: lily.fitzgibbon@sheffield.ac.uk

One of the hallmarks of human cognition is cognitive flexibility, the ability to adapt thoughts and behaviors according to changing task demands. Previous research has suggested that the number of different exemplars that must be processed within a task (the set size) can influence an individual's ability to switch flexibly between different tasks. This paper provides evidence that when tasks have a small set size, children's cognitive flexibility is impaired compared to when tasks have a large set size. This paper also offers insights into the mechanism by which this effect comes about. Understanding how set size interacts with task-switching informs the debate regarding the relative contributions of bottom-up priming and top-down control processes in the development of cognitive flexibility. We tested two accounts for the relationship between set size and cognitive flexibility: the (bottom-up) Stimulus-Task Priming account and the (top-down) Rule Representation account. Our findings offered support for the Stimulus-Task Priming account, but not for the Rule Representation account. They suggest that children are susceptible to bottom-up priming caused by stimulus repetition, and that this priming can impair their ability to switch between tasks. These findings make important theoretical and practical contributions to the executive function literature: theoretically, they show that the basic features of a task exert a significant influence on children's ability to flexibly shift between tasks through bottom-up priming effects. Practically, they suggest that children's cognitive flexibility may have been underestimated relative to adults', as paradigms used with children typically have a smaller set size than those used with adults. These findings also have applications in education, where they have the potential to inform teaching in key areas where cognitive flexibility is required, such as mathematics and literacy.

Keywords: cognitive flexibility, development, priming, executive function, set size, rule representation

INTRODUCTION

One of the hallmarks of human cognition is its flexibility. People are capable of flexibly adapting their thoughts and behaviors according to novel or changing environmental demands or task goals. For example, when switching between a Mac and a PC, different responses are often required to achieve the same goal, such as pressing a button in the top-left or top-right corner to close a browser window. Cognitive flexibility in adults and children is affected by the set size of the tasks involved—that is, the size of the pool of different stimuli that must be processed in the task (Kray and Eppinger, 2006; Kray et al., 2012). When a large pool of stimuli are used (a large set size), performance is better than when a small pool of stimuli are used (a small set size).

Set size is of theoretical importance because it informs the debate regarding the roles of top-down cognitive control and bottom-up priming in the development of cognitive flexibility (Cepeda et al., 2001). Set size is also methodologically important because one of the crucial differences between cognitive flexibility paradigms used with adults and those used with children is their set size (Cragg and Chevalier, 2012). Cognitive flexibility paradigms used with young children typically use a small number of stimuli (e.g., the Dimensional Change Card Sort, DCCS, Zelazo, 2006; and Shape School, Espy, 1997). In contrast, paradigms used with adults typically use a much larger set size (Rogers and Monsell, 1995; Richter and Yeung, 2012). Understanding the influence of set size in cognitive flexibility development can also better inform early school education for children in key areas such as mathematics and literacy where cognitive flexibility plays a central role (Bull and Scerif, 2001; St Clair-Thompson and Gathercole, 2006; Blair and Razza, 2007; Bull et al., 2008; Yeniad et al., 2013). This paper explores what role set size plays in cognitive flexibility during the early school years. We will first describe the development of cognitive flexibility during the early school years, and then discuss possible explanations for the role that set size might play in children's ability to switch flexibly between tasks.

When studying children's cognitive flexibility, researchers typically use paradigms that involve switching between two simple tasks, such as matching stimuli by their color and matching stimuli by their shape (Zelazo, 2006). By 3 years, children can perform either task well on its own, but typically fail to switch from one to the other (Zelazo et al., 1996). By 4 years, children can reliably make a single switch from one task to another (Zelazo et al., 1996) but experience difficulty switching flexibly back and forth

between the two tasks (Carlson, 2005; Hongwanishkul et al., 2005; Moriguchi and Hiraki, 2013). Around the age of five, children become able to flexibly switch back and forth between simple tasks (Chevalier and Blaye, 2009). At this age, response time starts to be a reliable metric of children's cognitive flexibility. This allows nuanced questions about the component processes required for cognitive flexibility to be investigated (Best and Miller, 2010). From around this age children are more likely to respond more slowly and less accurately when asked to switch from one task to another (i.e., on switch trials) than when asked to repeat the same task (i.e., on non-switch trials) (Dauvier et al., 2012). This is known as the switch cost. Switch costs tend to decrease with age (Crone et al., 2006; Huizinga et al., 2006; Chevalier and Blaye, 2009; Cragg and Nation, 2009; however this developmental trend is not reliable after the preschool years: see Dibbets and Jolles, 2006; Karbach and Kray, 2007; Kray et al., 2012). Switch costs do not diminish completely. They can reliably be found when adults must switch between tasks (for a review see Kiesel et al., 2010).

Examining switch costs in young children allows us to address important questions, such as what processes contribute to switch costs, and how these processes change during development (Best and Miller, 2010). Lessons drawn from adult participants suggest that switch costs reflect two distinct types of cognitive process. First, top-down control processes contribute to switch costs. These include the retrieval of task rules and the deliberate shifting of attention toward task-relevant stimulus attributes which are required on switch trials, but not on non-switch trials (Rogers and Monsell, 1995; Meiran, 1996; Monsell, 2003). Second, bottom-up priming processes contribute to switch costs. These include the priming of associations between stimuli and responses that build up over successive trials and facilitate non-switch trials but are detrimental to switch trials (Allport and Wylie, 2000; Waszak et al., 2003). It is not yet well understood to what extent each of these processes contributes to switch costs in children (see Cragg and Chevalier, 2012 for a review).

In this paper we explore the role of set size in the development of cognitive flexibility in children aged between 4 and 12 years. In particular, through manipulations of set size, we investigate the relative roles of top-down rule representation and bottom-up stimulus-task priming on cognitive flexibility during the early school years. The following sections explain how rule representation and stimulus-task priming relate to set size and the development of cognitive flexibility.

One mechanism by which set size would likely affect cognitive flexibility is through the way that task rules are represented. Consider a paradigm where children must switch between matching stimuli by their colors and matching them by their shapes. With a small set size [when the only colors in the task are (blue) and (red)], task rules can be efficiently represented in stimulus-specific terms—for example, "red blocks go in the red box" and "blue blocks go in the blue box." However, with a large set size, (for example, when there are many colors), it would be highly inefficient to formulate one rule for each individual color. It would be far more efficient to represent the task rules in abstract, dimension-level terms, such as "put the blocks into boxes that are the same color." Indeed, large pools of stimuli have been found to promote abstract categorization in toddlers (Perry et al., 2010). It is thus plausible that large and small set sizes create quite different task representations: a large set size on a task is likely to engender more abstract, dimension-level representations of task rules, whereas a small set size may engender more stimulus-specific representations of task rules.

Relevant to the relationship between set size and cognitive flexibility, evidence suggests that the way rules are represented determines how flexibly they can be switched between. It has been suggested that the early development of the prefrontal cortex supports abstract representation of task rules (Munakata et al., 2011) and cognitive flexibility (Moriguchi and Hiraki, 2013). Changes in the way that task rules are represented, from stimulus-specific representation to dimension-level representation, leads to better cognitive flexibility performance (Kharitonova et al., 2009; Kharitonova and Munakata, 2011). Thus, it is plausible that a large set size facilitates cognitive flexibility, by engendering dimension-level representation of task rules, and that smaller set sizes hinder rule switching, by engendering stimulus-specific rule representation.

The second mechanism by which set size would likely affect cognitive flexibility is through stimulus-task priming. Stimulus-task priming refers to the bottom-up process by which prior experience on a task leads to pairings that have been previously experienced being preferentially activated on later trials, regardless of whether they are currently task-relevant or not (Reuss et al., 2011). When the set size is small, individual stimuli appear more frequently—so there is more stimulus repetition—than when the set size is large. Associations between specific stimuli and specific tasks are thus more likely to build up with small set sizes than with large set sizes. Stimulus repetition has been shown to cause stimulus-task priming in adults, which contributes to greater switch costs (Waszak et al., 2003; Koch and Allport, 2006). Stimulus repetition is also detrimental to cognitive flexibility in preschool children (Müller et al., 2006; Experiment 3).

There are two reasons for thinking that stimulus-task priming might inflate switch costs. Firstly, on non-switch trials where the task repeats, if the stimulus was already associated with that task, then responses should be faster and more accurate because the currently relevant task was primed by the stimulus. Secondly, on switch trials, if the stimulus was associated with the previous (but no longer relevant) task, then responses should be slower and less accurate because the incorrect task was primed by the stimulus (Waszak and Hommel, 2007). Indeed, in a voluntary switching paradigm, where participants could choose to respond with a task repetition or a task switch on each trial, stimulus repetition was found to bias a task repetition response and stimulus change was found to bias a task switch response (Mayr and Bell, 2006).

To our knowledge only two studies have directly explored the effects of set size on cognitive flexibility in a developmental context (Kray and Eppinger, 2006; Kray et al., 2012). The first of these studies compared cognitive flexibility in young ($M = 21$ years) and older adults ($M = 66$ years) on a large and a small set size (Kray and Eppinger, 2006). It was found that the small set size induced greater switch costs than the large set size, but this effect was only seen in the older adults.

The second study (Kray et al., 2012) assessed cognitive flexibility in two groups of children (4- to 6-year-olds and 7- to

10-year-olds) and one group of young adults. Set size affected cognitive flexibility in two ways. First, when the set size was large, older children were better able to ignore task-irrelevant information than when set size was small. This effect was not seen for younger children or adults. Second, there was an effect of set size on conflict adaptation in older children. In some trials, both tasks would lead to the same response (compatible trials) so there was no conflict. In other trials, the relevant task would lead to one response and the non-relevant task would lead to the other response (incompatible trials) so there was conflict between the two tasks. The older children made greater control adjustments following incompatible trials in the small set size condition than the large set size condition. This suggests that the conflict that occurs between tasks is more salient with small set sizes than with large set sizes, and consequently results in better adjustment of control processes following its occurrence.

The absence of set size effects in younger children was surprising given the known developmental trends described above in both abstract rule representation (Munakata et al., 2012) and stimulus-task priming (Hommel et al., 2011). Furthermore, there is indirect evidence to suggest that preschool children's cognitive flexibility may be enhanced by increasing the task set size. For example, preschool children's ability to switch tasks has been improved by increasing the set size on the DCCS from two colors and two shapes to four colors and four shapes (Fisher, 2011a). However, the number of response options was also increased from two to four in that experiment, so it is unclear which of the two methodological changes was responsible for the facilitative effect. The absence of a set size effect in the youngest children in Kray et al.'s (2012) study may have been due to what the authors describe as the high "general demands on cognitive control processing" that the experimental paradigm entailed (Kray et al., 2012, p. 127). These high demands and the length of the test period may have also resulted in a high exclusion rate for the youngest age group (35%). Clearly, set size influences cognitive flexibility. However, the somewhat ambiguous findings indicate that further investigation is necessary.

Perhaps the most surprising finding was that increasing the set size did not reduce children's switching costs. This stands counter to the predictions drawn from the broader literature, and indeed counter to Kray et al.'s (2012) own predictions. However, this surprising absence of set size effect may in part have been due to the paradigm used. In one task, children categorized pictures as "animals" or "objects." This differs from typical developmental cognitive flexibility paradigms, which tend to be based on perceptual rather than conceptual features of the stimuli (FIST, Jacques and Zelazo, 2001; DCCS, Zelazo, 2006). Children are typically required to judge the color or shape of a stimulus, rather than its conceptual category membership. Conceptual categorization is a perfectly legitimate construct to study, though its use in cognitive flexibility paradigms may have attenuated the effects of set size in Kray et al.'s (2012) study for two reasons. First, children's perceptual processing is more robust than their conceptual processing (Fisher, 2011b), which would likely lead to stronger stimulus-task priming for perceptual than conceptual features. Second, there is also evidence to suggest that priming can occur at the level of semantic category as well as for individual stimuli (Waszak et al., 2004). Semantic category-level priming may have attenuated the facilitative effects of a large set size on switching costs.

The two experiments in this paper use the Switching Inhibition and Flexibility Task (SwIFT, Carroll and Cragg, 2012). This is a developmentally appropriate measure of cognitive flexibility that requires children to match stimuli according to their color or their shape. This kind of perceptual processing is known to be robust in young children (Zelazo et al., 1996; Fisher, 2011b). Processing demands that are orthogonal to cognitive flexibility are minimized: the goal setting demands are minimal because the task is cued on every trial with a transparent auditory cue (the word "color" or "shape"—Chevalier and Blaye, 2009). The working memory demands are low because the responses are intuitive so there is no need to maintain the appropriate responses for each task. The SwIFT thus gives a relatively pure measure of the costs of switching from one task to another by eliminating orthogonal processing costs that are present in other cognitive flexibility paradigms.

A review of the cognitive flexibility literature in general would lead us to predict a clear set-size effect on switch costs, although the few direct tests are tantalizingly inconclusive. On the basis of previous findings we would predict that a large set size would lead to smaller switch costs, and that a small set size would lead to larger switch costs. We expected that this effect would be largest in the youngest children, and would diminish with age. There were two reasons for this prediction: first, young children are less likely to spontaneously represent task rules in abstract, dimension level terms than older children (Kharitonova et al., 2009) and so would benefit most from a manipulation that engendered this type of rule representation. Second, children are more susceptible to stimulus-task priming than adults (Hommel et al., 2011), and if this relationship is linear, then younger children would be expected to be more affected by stimulus repetition than older children. We also expect to see a reduction in switch costs more generally during the early school years in line with findings from the Advanced DCCS paradigm (Chevalier and Blaye, 2009).

EXPERIMENT 1

METHODS

Participants

One hundred and forty nine 4- to 11-year-old children (80 female) were randomly selected from a larger sample attending Summer Scientist Week, a science engagement event at the University of Nottingham. Children were randomly allocated to one of two conditions: large set size or small set size. Each condition was further subdivided by age to give three similarly sized groups: 49 in the youngest age group (4;0- to 6;6-year-olds, $M = 5;2$, $SD = 0;8$, 27 females); 50 in the middle age group (6;7- to 8;4-year-olds, $M = 7;4$, $SD = 0;6$, 24 females); and 50 in the oldest age group (8;5- to 11;9-year-olds, $M = 9;10$, $SD = 1;0$, 29 females). Six further children were excluded because of missing data. Participants had no reported developmental disorders or special educational needs. Children's standardized scores on the British Picture Vocabulary Scale (BPVS) did not differ between the two test conditions [small set size: $M = 109.29$, large set size: $M = 109.73$, $t_{(138)} = -0.18$, $p = 0.86$]. This was indicative of similar levels of general cognitive functioning in the two groups.

BPVS scores were missing from nine participants. All the children were tested individually in a quiet room. Parental consent for participation in this research was obtained for all participants. The experimental procedure was approved by the School of Psychology Ethics Committee at the University of Nottingham.

Materials

Stimuli were presented and responses recorded on an Iiyama ProLite touch screen monitor connected to a Samsung P510 PC laptop running PsychoPy software (Peirce, 2007). Children responded by touching the relevant part of the screen and the program recorded their responses. The stimuli used in this task were nine simple novel shape outlines filled in with solid colors. The nine colors were all of equal saturation and brightness, and their hues were evenly distributed on a color wheel. Each image was approximately 6 × 8 cm.

Procedure

Children played a simple matching game. They were shown a prompt stimulus on a touchscreen computer, followed by two response stimuli. The prompt stimulus always had two dimensions (color and shape). On each trial, one response stimulus matched the prompt's color, and the other response stimulus matched the prompt's shape. Children were told to select the response stimulus that matched the prompt on the dimension relevant to the task for that trial (always either color or shape).

Trials were presented in two phases: the rule-learning phase and the task-switching phase. In the rule-learning phase there were two pure blocks of trials (6 trials each), in which children performed the same task for every trial. In one pure block children were required to match the stimuli by color, and in the other pure block they were required to match the stimuli by shape. The order of the pure blocks was counterbalanced between participants. In the task-switching phase there were three mixed blocks. During the mixed blocks (12 trials each) children were required to switch between the two tasks. The order of trials was pseudo-randomized such that some trials required children to perform the same task as the trial before (non-switch trials) and others required children to perform a different task to the trial before (switch trials). The first trial of each block was neither a non-switch nor a switch trial. There were a total of 16 non-switch trials and 17 switch trials. The number of trials was chosen to be developmentally appropriate for the younger participants, and was in line with previous research with 4- to 6-year-old children on the Shape School and Advanced DCCS paradigms (Zelazo, 2006; Chevalier et al., 2010; Blaye and Chevalier, 2011).

Before each pure block, children were shown an example trial using the standard task array of a prompt and two response stimuli. The task rules and the correct way to respond (touching the appropriate response stimulus with the index finger) were explained by the experimenter. The first pure block was also preceded by two practice trials. Feedback for correct and incorrect performance was given after each practice trial in the form of on-screen text and verbal feedback from the experimenter. For all experimental trials no feedback was given. If any practice trials were completed incorrectly two further practice trials were presented.

A graphical representation of the trial procedure can be found in **Figure 1**. Each trial began with a white screen showing a black outline of a rectangle (the prompt box) located at the top center of the screen. After a delay of 1000 ms, the prompt stimulus appeared in the prompt box, together with an auditory cue (a female voice saying "color" or "shape," as appropriate for that trial). After a further delay of 500 ms, two response stimuli appeared in the bottom left and right corners of the screen. One response stimulus matched the prompt on the color dimension, and the other response stimulus matched the prompt on the shape dimension. Neither response stimulus ever matched the prompt stimulus on both dimensions. All stimuli remained on the screen until children responded. Testing lasted approximately 15 min.

The experiment used a between-participants design, and there were two conditions, differing only in terms of set size. In the small set size condition there were two exemplars of color and two exemplars of shape (meaning that there were four stimuli in total). As in the Advanced DCCS, the target stimuli were kept constant, although their positions were counterbalanced. Within each block, each prompt stimulus was displayed six times, an equal number of times on color and shape trials, and approximately an equal number of times on non-switch and switch trials. There was an average of one intervening trial between recurrences of the same prompt stimulus.

In the large set size condition there were nine exemplars of color and nine exemplars of shape (meaning that there were 81 bidimensional stimuli in total). Stimulus selection was constrained so that a prompt stimulus never appeared more than once within a block, and the color and shape values never repeated on consecutive trials. On approximately half of the trials, one of the dimension values (the color or the shape) had occurred previously within the block. There were on average

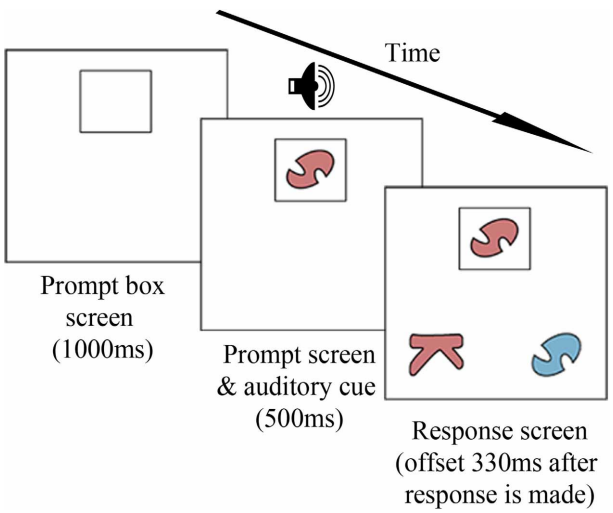


FIGURE 1 | Trial procedure. The prompt stimulus (**top**) matches one response stimulus (bottom **left**) according to its color and the other response stimulus (bottom **right**) according to its shape. An auditory cue ("color" or "shape") is onset concurrently with the prompt stimulus.

four intervening trials between recurrences of a dimension value within each block.

RESULTS

All analyses were performed after excluding the first trial from each block, since these trials were neither switch nor non-switch trials. Trials with RTs less than 200 ms or greater than 10,000 ms, and trials that were 2.5 standard deviations away from the individual's mean RT for that type of trial (5.4%), were excluded from the response time analysis. The response time analysis also included only correct trials that also followed a correct trial. This is because only these trials can be definitively classified as a non-switch trial or a switch trial. The mean number of trials entered into the analysis did not differ between the set size conditions, nor between switch and non-switch trials. Younger children contributed fewer trials to the analysis than older children because of higher error rates (Ms = 23.4, 25.4, and 27.4 trials entered for the youngest, middle and oldest age groups respectively). A natural logarithmic transformation was applied to the response time data to control for baseline changes in response time with age (Meiran, 1996; Chevalier et al., 2010). For clarity, untransformed RTs are presented in figures and tables.

Analyses were conducted separately for each of two dependent variables: mean accuracy and mean log-transformed RTs. To assess switch costs, mixed-measures ANOVAs were performed with trial type (non-switch vs. switch) as a within-participants variable, and age group (youngest vs. middle vs. oldest) and set size (small vs. large) as between-participants variables. We chose to use trial type (non-switch and switch) as a within-subjects variable rather than the difference score because this controls for both overall performance and processing speed differences between the experimental groups.

The analysis of accuracy data revealed a main effect of age, $F_{(2,\ 143)} = 10.24$, $p < 0.001$, $\eta^2 = 0.13$. Bonferroni-adjusted *post-hoc* tests revealed that the youngest group was less accurate than both the middle and oldest age groups (ps < 0.05). The middle age group was also less accurate than the oldest age group ($p < 0.05$). Analysis of trial type revealed a significant switch cost, $F_{(1,\ 143)} = 50.47$, $p < 0.001$, $\eta^2 = 0.26$, with less accurate performance on switch trials than non-switch trials (see **Table 1** for a summary of the means). Accuracy switch costs were not significantly different between age groups. There was no effect of set size on the accuracy of performance. This indicates that there were no baseline differences in overall accuracy on the tasks between the set size conditions.

The analysis of RT data revealed a main effect of age, $F_{(2,\ 143)} = 24.36$, $p < 0.001$, $\eta^2 = 0.25$. Bonferroni-adjusted *post-hoc* tests revealed that the youngest group was slower than both the middle and oldest age groups (ps < 0.01). The middle age group was also slower than the oldest age group ($p < 0.01$). Analysis of trial type revealed a significant switch cost, $F_{(1,\ 143)} = 22.57$, $p < 0.001$, $\eta^2 = 0.14$, with slower performance on switch trials than on non-switch trials. RT switch costs were not significantly different between age groups (see **Table 1** for a summary of the means).

There was an interaction between set size and trial type, $F_{(1,\ 143)} = 5.84$, $p < 0.05$, $\eta^2 = 0.04$, such that switch costs were greater in the small set size condition than in the large set size condition (see **Figure 2**). To investigate whether set size affected response times on switch trials, on non-switch trials, or on both, separate Bonferroni-adjusted *post-hoc* tests were performed for RTs on Switch trials and non-switch trials, comparing the small and large set-size conditions. Descriptively, non-switch trials were faster in the small set size condition than the large set size condition (see **Figure 2**). Conversely, switch trials were slower in the small set size condition than the large set size condition. However, these differences were not significant (ps > 0.1). There was no overall effect of set size on RT, which indicates that the set size conditions did not differ in terms of processing speed.

Paired samples t-tests were conducted to determine whether RTs on switch and non-switch trials were significantly different for each set size condition (i.e., whether there was a significant switch cost). In the small set size condition, switch trials were significantly slower than non-switch trials, $t_{(69)} = 4.57$, $p < 0.01$.

Table 1 | Mean accuracy rates and response times by trial type in Experiment 1.

Set size	*N*	Age in years	Female	Accuracy (%)				Response times (ms)			
				Pure blocks	Non-switch	Switch	Switch cost	Pure blocks	Non-switch	Switch	Switch cost
YOUNGEST CHILDREN											
Small	25	5;2 (0;8)	18	96.3 (5.9)	89.5 (13.2)	78.4 (16.9)	11.1 (12.3)	1482 (440)	1578 (365)	1916 (715)	337 (584)
Large	24	5;3 (0;9)	9	96.9 (4.1)	89.6 (11.0)	84.1 (12.8)	5.5 (12.1)	1362 (287)	1714 (382)	1773 (390)	59 (203)
MIDDLE CHILDREN											
Small	23	7;4 (0;7)	11	96.0 (5.5)	93.5 (7.9)	85.2 (7.1)	8.3 (10.2)	1104 (300)	1391 (391)	1493 (490)	102 (183)
Large	27	7;5 (0;6)	13	96.3 (4.8)	93.1 (8.6)	86.3 (12.2)	6.8 (10.6)	1186 (293)	1484 (364)	1571 (434)	87 (235)
OLDEST CHILDREN											
Small	22	9;9 (1;1)	12	97.0 (6.1)	94.6 (5.9)	91.2 (13.1)	3.4 (14.5)	914 (192)	1168 (318)	1258 (450)	91 (219)
Large	28	9;8 (1;0)	17	96.4 (5.8)	96.4 (5.2)	92.0 (7.2)	4.4 (8.2)	999 (293)	1239 (257)	1272 (310)	33 (180)

For ease of reference, switch costs are also shown. Standard deviations are presented in parentheses.

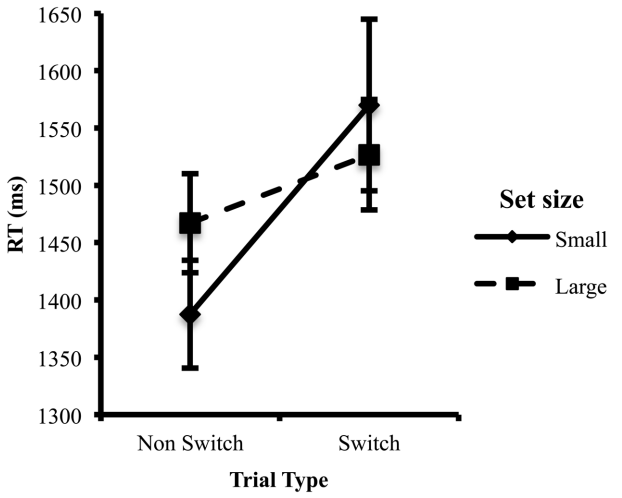


FIGURE 2 | Mean RT in the switch and non-switch trials as a function of the set size. Error bars indicate the standard error of the mean.

In the large set size condition, switch trials were marginally slower than non-switch trials, $t_{(78)} = 1.91, p = 0.060$.

To investigate whether there were more gradual age-related changes in cognitive flexibility in our sample, accuracy and RT switch costs (difference scores between switch and non-switch trials) were entered into a bivariate correlation analysis with age (mean-centered). Both types of switch cost were negatively correlated with age [accuracy: $r_{(147)} = -0.18, p = 0.033$; RT: $r_{(147)} = -0.17, p = 0.037$].

DISCUSSION

In Experiment 1, a switching paradigm with minimal incidental demands was used to investigate the effects of set size on cognitive flexibility during the early school years. The results showed significant age-related changes in accuracy and RT on both non-switch and switch trials. However, contrary to our predictions there were no significant differences in switch costs between the age groups. This is consistent with other studies that have found little or no change in switch costs over the early school period (Dibbets and Jolles, 2006; Karbach and Kray, 2007; Kray et al., 2012; though see also Chevalier and Blaye, 2009; Cragg and Nation, 2009). However, correlational analyses did reveal a developmental trend of reduced switch costs with age for both the accuracy and the RT data. This suggests that there are developmental changes to switch costs, but that these are gradual and may be harder to detect when data are categorized for analysis by age group. This gradual trend may explain the inconsistency in the literature with regards to age-related changes in switch costs.

In line with our predictions, when the set size was small, response-time switch costs were greater than when the set size was large. This indicates that a smaller set size leads to more interference between tasks, and is consistent with Kray et al.'s (2012) findings. It is also consistent with findings that show switch costs to be greater when there is more stimulus overlap between tasks (Waszak et al., 2003; Koch and Allport, 2006). These data clearly demonstrate, then, that a set size effect is apparent in children's cognitive flexibility. Furthermore, the difference between switch and non-switch trials was only marginally significant when the set size was large. This shows that the cognitive processes that contribute to switch costs are affected by the set size. Thus, to understand the processes that contribute to switch costs in children, it is necessary to understand what drives the set-size effect found in Experiment 1. The mechanisms that underpin this effect remain to be elucidated.

Experiment 1 identified that increasing the set size of the tasks reduced the cost of switching between them. We have identified two candidate cognitive mechanisms that could explain why increasing the set size reduces switch costs. The first explanation, which we refer to as the Rule Representation account, posits a mechanism in which set size affects the way that task rules are represented, which influences cognitive flexibility (Kharitonova et al., 2009). The larger set size may engender abstract, dimension-level representations of the task rules, whereas the smaller set size may engender stimulus-specific representations of the task rules. More abstract representations should lead to more flexible switching between tasks (Kharitonova et al., 2009).

The second explanation, which we refer to as the Stimulus-Task Priming account, posits a mechanism in which set size affects bottom-up priming of stimulus-to-task mappings (Waszak et al., 2003; Koch and Allport, 2006). This is because when the set size is small, individual stimuli repeat more often, both within and between the two tasks. This would be expected to lead to associations between stimuli and tasks which would both facilitate repeating a task from one trial to the next, and impair switching between tasks. In contrast, when the set size is large, individual stimuli repeat less often. This would be expected to lead to much less pronounced priming effects relative to a smaller set size. The direction of the means in Experiment 1 are consistent with this account insofar as non-switch trials were slower when the set size was large than when it was small, and switch trials were faster when the set size was large than when it was small. In adults, associations between stimuli and tasks can lead to switch costs even after 100 intervening trials (Waszak et al., 2003). The duration of stimulus-task associations in children is not yet known. The effect may last as long as it does in adults, or it may be limited to consecutive trials. To maximize the chances of detecting a stimulus-task priming effect—if it exists—the set size conditions differed both in terms of frequency of stimulus repetitions, and in terms of the number of intervening trials between repetitions. The small set size condition had many trials where stimuli repeated from one trial to the next. The large set size condition had no trials where stimuli repeated on consecutive trials. However, within each condition, the number of intervening trials was not systematically varied, so the duration of stimulus-task priming was not investigated here.

To test the Rule Representation account and the Stimulus-Task Priming account directly, it is necessary to tease apart two things: the rule representation effects that are initiated during the rule-learning phase, and the priming effects that occur during the task-switching phase. To do this, set size was varied independently in the rule-learning phase and in the task-switching phase.

Set size was either large or small. This two-by-two design yielded four conditions, differing firstly according to the set size in the rule-learning phase, and secondly according to the set size in the later task-switching phase. Note that the Rule Representation and Stimulus-Task Priming accounts are not mutually exclusive. The set size effect observed in Experiment 1 may be best explained by one process, or by the other, or by both together. Experiment 2 seeks to address this question.

The Rule Representation account predicts that the set size in the rule-learning phase would have an effect on switch costs. Specifically, if the set size was large in the rule-learning phase, then children would be more likely to form more flexible dimension-level rule representations, and if the set size was small in the rule learning phase then children would be more likely to form less flexible stimulus-specific rule representations. Thus, a large set size during the rule-learning phase should lead to smaller switch costs than a small set size during the rule-learning phase. The Stimulus-Task Priming account predicts that the set size in the task-switching phase would have an effect on switch costs. Specifically, when the set size was small in the task-switching phase, then there would be more stimulus repetition between tasks than when the set size was large in the task-switching phase. This would lead to larger stimulus-task priming effects when the set size was small in the task-switching phase than when the set size was large in the task-switching phase. Thus, a large set size during the task-switching phase should lead to smaller switch costs than a small set size during the task-switching phase. Note that children form representations of the task rules quickly. Even at 3 years of age, children are capable of forming abstract representations of task rules after as few as six trials (Kharitonova et al., 2009).

Experiment 2 also explores developmental changes in the roles of rule representation and stimulus-task priming on cognitive flexibility with development. With age, children become more likely to spontaneously use abstract rule representations (Kharitonova et al., 2009; Kharitonova and Munakata, 2011). It was thus expected that having a large set size in the rule-learning phase would have the greatest facilitative effect on the youngest children in the sample, since they would be the least likely to spontaneously form abstract rule representations for the small set size condition, and thus the most likely to benefit from a condition that engenders it. Similarly, stimulus-task associations have been shown to be more robust in children than in adults (Hommel et al., 2011). If the strength of these associations follows a linear relationship through development, the effects of set size during the task-switching phase should also be strongest for the youngest children in our sample.

EXPERIMENT 2

METHODS

Participants

Two hundred and forty three 5- to 10-year-old children (128 female) from two suburban primary schools in the UK took part: 84 in the youngest age group (5;3- to 6; 6-year-olds, $M = 6;0$, $SD = 0$; 4, 46 females), 79 in the middle age group (7;2- to 8;6-year-olds, $M = 8;0$, $SD = 0;4$, 39 females) and 80 in the oldest age group (9;5- to 10;6-year-olds, $M = 10;0$, $SD = 0;4$, 42 females). Six further children were excluded either because they failed to follow instructions ($N = 3$) or because of missing data ($N = 3$). Participants had no known developmental disorders or special educational needs. Participants were randomly assigned to one of the four conditions. All the children were tested individually in a quiet room in their schools. Parental consent for participation in this research was obtained for all participants. The experimental procedure was approved by the Department of Psychology Ethics Committee at the University of Sheffield.

Materials and procedure

The stimuli and materials for Experiment 2 were the same as for Experiment 1, except that a Dell T7570 laptop running E-Prime v1.2 (Psychological Software Tools, Pittsburgh, PA) software was connected to an ATM-152ROHACB2D touch screen monitor. Set size was varied using a 2×2 design with two levels of set size (small and large) varying across the two phases of the experiment (the initial rule-learning phase, and the subsequent task-switching phase). This resulted in four conditions: the small:small condition had a small set size in the rule-learning phase and a small set size in the task-switching phase; the small:large condition had a small set size in the rule-learning phase and a large set size in the task-switching phase; the large:small condition had a large set size in the rule-learning phase and a small set size in the task-switching phase; and the large:large condition had a large set size in the rule-learning phase and a large set size in the task-switching phase. (Note that the small:small condition and the large:large condition were identical to the small and large set size conditions in Experiment 1.) Stimulus selection was constrained in the same way as in Experiment 1.

RESULTS

The mean accuracy and mean log-transformed response time (RT) were calculated for each child for each trial type. Trials with RT less than 200 ms or greater than 10,000 ms, and trials 2.5 standard deviations or greater away from the individual's mean RT for that type of trial (5.0%) were excluded from the response-time analysis. Response-time analyses included only correct trials that followed a correct trial. The mean number of trials entered into the analysis did not differ between the set size conditions, nor between switch and non-switch trials. Younger children contributed fewer trials to the analysis than older children because of higher error rates (Ms = 23.1, 25.5, and 24.3 trials entered for the youngest, middle and oldest age groups respectively).

All analyses were performed separately for each dependent variable of interest: mean accuracy and mean log-transformed RTs. To assess switch costs, a mixed measures ANOVA was performed with trial type (non-switch vs. switch) as a within-participants factor, and age (youngest vs. middle vs. oldest), set size in the rule-learning phase (small vs. large) and set size in the task-switching phase (small vs. large) as between-participant factors.

The analysis of accuracy data revealed a main effect of age, $F_{(2,\ 231)} = 15.72$, $p < 0.001$, $\eta^2 = 0.0.12$. Bonferroni-adjusted *post-hoc* tests showed that the youngest children were less accurate than both the middle children and the oldest children ($ps <$ 0.001; see **Table 2** for a summary of the means). There was no

Table 2 | Mean accuracy rates and response times by trial type in Experiment 2.

Set size	*N*	Age in months	Female	Accuracy (%)				Response times (ms)			
				Pure blocks	Non-switch	Switch	Switch cost	Pure blocks	Non-switch	Switch	Switch cost
YOUNGEST CHILDREN											
Small:Small	21	6;1 (0;4)	11	96.8 (5.6)	89.0 (14.9)	86.6 (17.7)	2.4 (12.3)	1339 (316)	1697 (427)	1785 (428)	88 (235)
Small:Large	19	6;1 (0;4)	12	93.8 (8.5)	89.8 (17.5)	84.2 (12.6)	5.6 (10.4)	1321 (389)	1679 (339)	1772 (458)	93 (306)
Large:Small	23	5;11 (0;4)	13	98.6 (4.1)	92.9 (7.6)	82.6 (14.6)	10.3 (11.1)	1540 (340)	1757 (424)	1871 (518)	115 (304)
Large:Large	21	6;0 (0;4)	10	95.2 (6.8)	96.4 (7.0)	88.5 (12.7)	7.9 (10.8)	1410 (387)	1761 (503)	1700 (387)	−62 (230)
MIDDLE CHILDREN											
Small:Small	20	7;11 (0;3)	9	98.3 (3.4)	98.4 (2.8)	93.8 (6.7)	4.6 (6.2)	1167 (385)	1312 (301)	1389 (345)	77 (215)
Small:Large	20	7;11 (0;4)	9	98.3 (4.4)	98.1 (4.1)	91.5 (9.0)	6.7 (8.4)	1170 (280)	1427 (425)	1529 (376)	102 (248)
Large:Small	20	7;11 (0;4)	11	97.9 (4.6)	96.3 (5.1)	93.2 (5.8)	3.0 (7.4)	1404 (556)	1355 (367)	1499 (342)	144 (173)
Large:Large	19	8;0 (0;4)	10	96.7 (6.8)	95.4 (5.8)	90.4 (9.2)	5.0 (9.7)	1261 (257)	1403 (321)	1426 (271)	23 (233)
OLDEST CHILDREN											
Small:Small	19	10;0 (0;3)	10	98.2 (4.5)	98.4 (4.6)	95.7 (7.6)	2.7 (4.9)	932 (162)	1140 (243)	1208 (223)	69 (164)
Small:Large	20	10;0 (0;3)	10	99.2 (3.7)	95.6 (5.0)	93.2 (6.7)	2.4 (4.5)	934 (171)	1184 (253)	1252 (243)	68 (145)
Large:Small	20	10;0 (0;4)	11	99.2 (2.6)	95.9 (5.1)	92.4 (6.1)	3.6 (6.5)	1076 (280)	1186 (401)	1293 (314)	107 (239)
Large:Large	21	10;0 (0;3)	11	97.6 (4.7)	97.6 (3.1)	94.1 (6.7)	3.5 (6.3)	1006 (139)	1213 (243)	1228 (207)	15 (196)

For ease of reference, switch costs are also shown. Standard deviations are presented in parentheses.

significant difference between the accuracy of the middle and oldest children. Analysis of trial type revealed a significant switch cost, $F_{(1,\ 231)} = 75.23$, $p < 0.001$, $\eta^2 = 0.24$, with less accurate performance on switch trials than non-switch trials. There was an interaction between age group and trial type, $F_{(2,\ 231)} = 3.41$, $p < 0.05$, $\eta^2 = 0.03$. Bonferroni-adjusted *post-hoc* tests showed that the youngest children had greater switch costs for accuracy than the oldest children ($p < 0.05$). No other age comparisons were significant. There was also a three-way interaction between age, trial type and set size in the rule-learning phase, $F_{(2,\ 231)} = 3.15$, $p < 0.05$, $\eta^2 = 0.03$. Bonferroni-adjusted *post-hoc* tests showed that for the youngest age group, switch costs were greater when the set size in the rule-learning phase was large than when it was small ($p < 0.05$). Set size in the rule-learning phase did not affect switch costs for the middle or oldest age groups ($ps > 0.1$). There was no effect of set size on the accuracy of performance. This indicates that there were no baseline differences in overall accuracy on the tasks between the set size conditions.

The analysis of RT data revealed a main effect of age, $F_{(2,\ 231)} = 53.78$, $p < 0.001$, $\eta^2 = 0.32$. Bonferroni-adjusted *post-hoc* tests revealed that the youngest group was slower than both the middle and oldest age groups ($ps < 0.01$). The middle age group was also slower than the oldest age group ($p < 0.01$). Analysis of trial type revealed a significant switch cost, $F_{(1,\ 231)} = 33.84$, $p < 0.001$, $\eta^2 = 0.13$, with slower performance on switch trials than non-switch trials (see **Table 2** for the mean transformed RTs).

Two further significant interactions were revealed in the RT analysis. First, the set size in the task-switching phase interacted with the trial type $F_{(1,\ 231)} = 7.03$, $p < 0.01$, $\eta^2 = 0.03$. Overall, switch costs were larger when the set size in the task-switching phase was small ($M = 100$ ms) than when the set size in the task-switching phase was large ($M = 39$ ms). Separate Bonferroni-adjusted *post-hoc* tests were performed for RTs on Switch and Non-switch trials comparing conditions with a small set size in the task-switching phase and those with a large set size in the task-switching phase. Descriptively, when there was a small set size in the task-switching phase, non-switch trials were faster than when there was a large set size in the task-switching phase. Conversely, when there was a small set size in the task-switching phase, switch trials were slower than when there was a large set size in the task-switching phase (see **Figure 3**). However, these differences were not significant ($ps > 0.1$).

Second, there was a three-way interaction between the set size in the rule-learning phase, the set size in the task-switching phase and the trial type, $F_{(1,\ 231)} = 7.73$, $p < 0.01$, $\eta^2 = 0.03$. Two pairwise Bonferroni-adjusted *post-hoc* tests were conducted. One compared switching costs between the two conditions with a large set size in the rule-learning phase (the large:small and large:large conditions). The other compared the two conditions with a small set size in the rule-learning phase (the small:small and small:large conditions). When the set size in the rule-learning phase was large, switch costs were affected by the set size in the task-switching phase ($p < 0.05$): switch costs were larger in the large:small condition ($M = 122$ ms) than in the large:large condition ($M = -9$ ms; this value is negative because switch trials were faster than non-switch trials in this condition). When the set size in the rule-learning phase was small, switch costs were not affected by the set size in the task-switching phase ($p > 0.1$, see **Figure 3**). There was no overall effect of set size on RT, which indicates that the groups did not differ in terms of processing speed.

Paired samples *t*-tests were conducted for each set size condition to determine whether RTs on switch and non-switch trials were significantly different (i.e., whether there was a significant switch cost). For the small:small, small:large and large:small conditions, switch trials were slower than non-switch trials, $ts > 1$, $ps < 0.01$. For the large:large condition, RTs on switch and non-switch trials did not differ significantly.

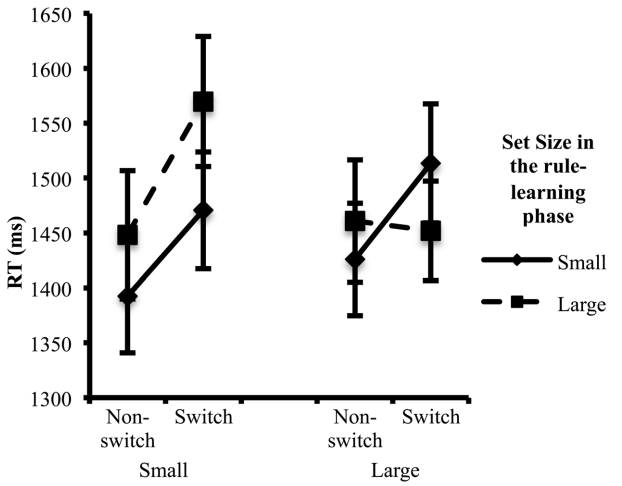


FIGURE 3 | Mean RT in the switch and non-switch trials as a function of the set size in the rule-learning and task-switching phases. Error bars indicate the standard error of the mean.

To test whether the set-size effect found in Experiment 1 was replicated in Experiment 2, planned comparisons of switch costs in the small:small and large:large conditions were conducted. These showed that there were greater switch costs in the small:small condition than in the large:large condition, $t_{(109)} = 2.30$, $p < 0.05$. Thus, the set-size effect reported in Experiment 1 was also replicated in Experiment 2.

Finally, to investigate whether there were more gradual age-related changes in cognitive flexibility in our sample, accuracy and RT switch costs (difference scores between switch and non-switch trials) were entered into a bivariate correlation analysis with age. Accuracy switch costs were negatively correlated with age (mean-centered), $r_{(241)} = -0.16$, $p = 0.011$. However, reaction time switch costs were not related to age: $r_{(241)} = -0.056$, $p = 0.38$.

DISCUSSION

Experiment 2 replicated the key finding of Experiment 1, namely that with school-aged children, a small set size leads to larger switch costs than a large set size. In addition, Experiment 2 extended our understanding of how this effect comes about by testing the two accounts of this effect: the Rule Representation account and the Stimulus-Task Priming account. There was no support for the Rule Representation account: a large set size in the rule-learning phase did not lead to reduced switch costs compared to a small set size. Conversely, there was support for the Stimulus-Task Priming account: a large set size in the task-switching phase led to reduced switch costs compared to a small set size in the task-switching phase. However, the presence of a three-way interaction between trial type, set size in the rule-learning phase and the set size in the task-switching phase suggests that there is a more complex story to be told. Further discussion of these accounts in light of the current findings can be found in the general discussion.

Contrary to previous research, Experiment 2 did not provide evidence for age-related changes in stimulus-task priming. Previous research has shown that 9- and 10-year-old children are more susceptible to stimulus-task priming than young adults (Hommel et al., 2011). Experiment 2 included children up to 10 years of age. The lack of interaction between age, trial type and the set size in the task-switching phase suggests that developmental changes in the ability to overcome stimulus task priming may be limited to later childhood and adolescence. An alternative explanation is that stimulus-task priming may be developmentally invariant. This view is consistent with research demonstrating that implicit learning processes such as priming of stimulus-response associations does not change from infancy to adulthood (Amso and Davidow, 2012).

GENERAL DISCUSSION

The two experiments presented in this paper shed important new light on how set size influences cognitive flexibility during development. They extend previous research by showing that children's switch costs can be reduced by increasing the set size. This finding is directly relevant to preschool research into cognitive flexibility. The most widely used paradigm, the DCCS (Zelazo et al., 1996; Müller et al., 2006), typically uses a small set size (usually comprising four stimuli in total). Experiment 1 and 2 indicate that using a small set size is likely to be particularly difficult for children, and that studies using such paradigms may systematically underestimate children's cognitive flexibility.

The present findings build on the work of Kray et al. (2012) who showed that when children switched between tasks, they experienced less interference when there was a large set size than when there was a small set size. The two experiments presented in this paper extend those findings to show that during the early school years, switch costs can also be reduced when the set size is increased. By manipulating set size in a paradigm that has minimal incidental demands and which uses simple perceptual categorization tasks, these experiments showed a robust effect of set size on switch costs for children from 4 to 12 years.

The Rule Representation account of the set size effect was not supported. According to this account, the set size affects the way task rules are represented which affects switch costs. A small set size during the rule-learning phase was expected to result in less flexible stimulus-specific rule representations while a large set size during the rule-learning phase was expected to result in more flexible dimension-level rule representation. None of the predictions derived from the Rule Representation account were borne out in the findings of Experiment 2. Indeed, directly contrary to the prediction of the Rule Representation account, the youngest children's accuracy switch costs were greater when there was a large set size in the rule-learning phase than when there was a small set size in the rule-learning phase. These findings suggest that the way that task rules are represented does not drive the facilitative effects of a large set size on switch costs.

Furthermore, the lack of effect of set size in the rule-learning phase on switch costs raises questions over the robustness of the association between abstraction and flexibility. Previous research has shown that children who form dimension-level representations of task rules have better cognitive flexibility than

children with stimulus-specific rule representations (Kharitonova et al., 2009; Kharitonova and Munakata, 2011). However, in Experiment 2, there was no main effect of set size in the rule-learning phase. Engendering dimension-level rule representation by presenting participants with a large set size in the rule-learning phase did not increase later flexibility during the task-switching phase. It is possible that the rule-learning phase was too short to engender stable rule representations that persisted into the task-switching phase. However, research with children as young as 3 years demonstrates that dimension level and stimulus-specific rule representations can be formed after six trials (Kharitonova et al., 2009). In the two experiments presented in this paper, even the youngest children made very few mistakes during the rule-learning phase. This suggests that the rules were intuitive and easy to learn. It remains a question for future research whether more trials during the rule-learning phase would lead to more persistent representations of task rules.

The Stimulus-Task Priming account is supported by the findings of the two experiments presented in this paper. According to this account, the set size in the task-switching phase affects the amount of stimulus-task priming that occurs which affects switch costs. More stimulus repetition occurs when the set size was small in the task-switching phase than when the set size was large in the task-switching phase. This was expected to result in more stimulus-task priming and so greater switch costs when the set size was small in the task-switching phase than when the set size was large in the task-switching phase. First, in both experiments, when the set size in the task-switching phase was small, non-switch trials were faster and switch trials were slower than when the set size in the task-switching phase was large (although these differences were not statistically significant). This was consistent with predictions from the Stimulus-Task Priming account, since stimulus-task priming should be facilitative for non-switch trials and detrimental for switch trials (Waszak and Hommel, 2007). Second, in Experiment 2 there was a main effect of set size in the task-switching phase. This provides evidence that the link between set size and cognitive flexibility is mediated by stimulus-task priming that occurs as a result of stimulus repetition during the task-switching phase. This bottom-up process includes priming as a result of stimulus repetition both within task (which facilitates task repetition) and between the two tasks (which impairs task switching). The findings of this study suggest that young children's cognitive flexibility is affected by stimulus-task priming and that this priming contributes to switch costs.

However, two findings from Experiment 2 suggest that switch costs cannot be solely driven by bottom-up processes. First, switch costs were found in the near-absence of stimulus-task priming, which suggests that top-down processes may also contribute to switch costs. Specifically, in Experiment 2, significant switch costs were found in the small:large condition. Recall that in this condition, very little stimulus repetition occurs during the task-switching phase. The switch costs that occur here are thus unlikely to be driven by stimulus-task priming. Second, there was a three-way interaction between trial type, set size in the rule-learning phase and set size in the task-switching phase. This shows that set size during the rule-learning phase has carryover effects that moderate the effects of set size in the task-switching phase on switch costs.

Both of the above findings can be explained by the same top-down mechanism. We suggest that exposure to a small set size in the rule-learning phase led children to expect high levels of conflict between the tasks in the task-switching phase. This would have promoted more engagement of top-down control processes. It is plausible that children will prepare more for task switches under conditions where they expect high levels of conflict. Thus, switch costs likely occur in the small:large condition as a result of greater engagement from top-down control processes on switch trials than non-switch trials. These same top-down control processes likely attenuate the effects of set size in the task switching phase. This explains the three-way interaction found in the RT analysis. This explanation is consistent with the findings of Kray et al. (2012), who found that children adapted better to conflict with a small set size than a large set size.

Together, these findings suggest that a combination of top-down and bottom-up processes contribute to switching costs in early school age children, and that these are differentially affected by manipulations of set size. This proposal is entirely consistent with findings from the adult literature which suggest that stimulus-task priming only partially accounts for the cost of switching from one task to another (for a review, see Vandierendonck et al., 2010). However, the developmental trajectory of this interaction is still uncertain. Although we expected to find a shift from bottom-up to top-down processing with development (Munakata et al., 2012), our findings did not wholly bear this out. We found little evidence for reliable developmental change in cognitive flexibility. It is true that accuracy switch costs were negatively correlated with age in both experiments. However, RT switch costs were negatively correlated with age in Experiment 1, but not related to age in Experiment 2. This inconsistency may in part have been due to the variability of performance in our youngest age group (see standard deviations in **Tables 1**, **2**). It is common for response times to be variable for children of this age (for example see Chevalier and Blaye, 2009). This variability may be exacerbated by heterogeneity in cognitive strategies employed by 5- and 6-year-olds when engaged in task switching (Dauvier et al., 2012).

Clearly, set size influences different facets of cognitive flexibility in conflicting ways. Having a smaller set size in the task-switching phase impaired cognitive flexibility (leading to larger RT switch costs) by increasing the amount of stimulus-task priming that subjects were subjected to. Conversely, having a smaller set size in the rule-learning phase directly enhanced cognitive flexibility for the youngest children (leading to smaller accuracy switch costs), and attenuated the effects of stimulus-task priming on response time switch costs by increasing engagement of top-down control processes. This highlights the complexity of cognitive flexibility, demonstrating that a multitude of processes must work in harmony to produce flexible behavior (Cragg and Nation, 2010; Ionescu, 2012). However, it also raises the interesting possibility that there are multiple routes by which cognitive flexibility can be influenced, and that set size may act in contrasting ways depending on the stage in the task.

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Blaye, A., and Chevalier, N. (2011). The role of goal representation in preschoolers' flexibility and inhibition. *J. Exp. Child Psychol.* 108, 469–483. doi: 10.1016/j.jecp.2010.09.006

Bull, R., Espy, K. A., and Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev. Neuropsychol.* 33, 205–228. doi: 10.1080/87565640801982312

Bull, R., and Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Dev. Neuropsychol.* 19, 273–293. doi: 10.1207/S15326942DN1903_3

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carroll, D. J., and Cragg, L. (2012). "Precursors to cognitive flexibility: rule-switching in two-year-olds," in *Paper presented at the Development of Executive Functions Workshop* (Utrecht, Netherlands).

Cepeda, N. J., Kramer, A. F., and Gonzalez De Sather, J. C. (2001). Changes in executive control across the life span: examination of task-switching performance. *Dev. Psychol.* 37, 715–730. doi: 10.1037/0012-1649.37.5.715

Chevalier, N., and Blaye, A. (2009). Setting goals to switch between tasks: effect of cue transparency on children's cognitive flexibility. *Dev. Psychol.* 45, 782–797. doi: 10.1037/a0015409

Chevalier, N., Blaye, A., Dufau, S., and Lucenet, J. (2010). What visual information do children and adults consider while switching between tasks? Eye-tracking investigation of cognitive flexibility development. *Dev. Psychol.* 46, 955–972. doi: 10.1037/a0019674

Cragg, L., and Chevalier, N. (2012). The processes underlying flexibility in childhood. *Q. J. Exp. Psychol.* 65, 209–232. doi: 10.1080/17470210903204618

Cragg, L., and Nation, K. (2009). Shifting development in mid-childhood: the influence of between-task interference. *Dev. Psychol.* 45, 1465–1479. doi: 10.1037/a0015360

Cragg, L., and Nation, K. (2010). Language and the development of cognitive control. *Top. Cogn. Sci.* 2, 631–642. doi: 10.1111/j.1756-8765.2009.01080.x

Crone, E. A., Donohue, S. E., Honomichl, R., Wendelken, C., and Bunge, S. A. (2006). Brain regions mediating flexible rule use during development. *J. Neurosci.* 26, 11239–11247. doi: 10.1523/JNEUROSCI.2165-06.2006

Dauvier, B., Chevalier, N., and Blaye, A. (2012). Using finite mixture of GLMs to explore variability in children's flexibility in a task-switching paradigm. *Cogn. Dev.* 27, 440–454. doi: 10.1016/j.cogdev.2012.07.004

Dibbets, P., and Jolles, J. (2006). The switch task for children: measuring mental flexibility in young children. *Cogn. Dev.* 21, 60–71. doi: 10.1016/j.cogdev.2005.09.004

Espy, K. A. (1997). The shape school: assessing executive function in preschool children. *Dev. Neuropsychol.* 13, 495–499. doi: 10.1080/87565649709540690

Fisher, A. V. (2011a). Automatic shifts of attention in the dimensional change card sort task: subtle changes in task materials lead to flexible switching. *J. Exp. Child Psychol.* 108, 211–219. doi: 10.1016/j.jecp.2010.07.001

Fisher, A. V. (2011b). Processing of perceptual information is more robust than processing of conceptual information in preschool-age children: evidence from costs of switching. *Cognition* 119, 253–264. doi: 10.1016/j.cognition.2011.01.015

Hommel, B., Kray, J., and Lindenberger, U. (2011). Feature integration across the lifespan: stickier stimulus–response bindings in children and older adults. *Front. Psychol.* 2:268. doi: 10.3389/fpsyg.2011.00268

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., and Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: age-related changes and individual differences. *Dev. Neuropsychol.* 28, 617–644. doi: 10.1207/s15326942dn2802_4

Hübner, M., Dreisbach, G., Haider, H., and Kluwe, R. H. (2003). Backward inhibition as a means of sequential task-set control: evidence for reduction of task competition. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 289–297. doi: 10.1037/0278-7393.29.2.289

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas Psychol.* 30, 190–200. doi: 10.1016/j.newideapsych.2011.11.001

Jacques, S., and Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): a measure of executive function in preschoolers. *Dev. Neuropsychol.* 20, 573–591. doi: 10.1207/S15326942DN2003_2

Karbach, J., and Kray, J. (2007). Developmental changes in switching between mental task sets: the influence of verbal labeling in childhood. *J. Cogn. Dev.* 8, 205–236. doi: 10.1080/15248370701202430

Kharitonova, M., Chien, S., Colunga, E., and Munakata, Y. (2009). More than a matter of getting 'unstuck': flexible thinkers use more abstract representations than perseverators. *Dev. Sci.* 12, 662–669. doi: 10.1111/j.1467-7687.2008.00799.x

Kharitonova, M., and Munakata, Y. (2011). The role of representations in executive function: investigating a developmental link between flexibility and abstraction. *Front. Psychol.* 2:347. doi: 10.3389/fpsyg.2011.00347

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., et al. (2010). Control and interference in task switching–a review. *Psychol. Bull.* 136, 849. doi: 10.1037/a0019842

Koch, I., and Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Mem. Cognit.* 34, 433–444. doi: 10.3758/BF03193420

Kray, J., and Eppinger, B. (2006). Effects of associative learning on age differences in task-set switching. *Acta Psychol.* 123, 187–203. doi: 10.1016/j.actpsy.2005.12.009

Kray, J., Karbach, J., and Blaye, A. (2012). The influence of stimulus-set size on developmental changes in cognitive control and conflict adaptation. *Acta Psychol.* 140, 119–128. doi: 10.1016/j.actpsy.2012.03.005

Mayr, U., and Bell, T. (2006). On how to be unpredictable evidence from the voluntary task-switching paradigm. *Psychol. Sci.* 17, 774–780. doi: 10.1111/j.1467-9280.2006.01781.x

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 22:1423. doi: 10.1037/0278-7393.22.6.1423

Monsell, S. (2003). Task switching. *Trends Cogn. Sci.* 7, 134–140. doi: 10.1016/S1364-6613(03)00028-7

Moriguchi, Y., and Hiraki, K. (2013). Behavioral and neural differences during two versions of cognitive shifting tasks in young children and adults. *Dev. Psychobiol.* doi: 10.1002/dev.21145/full. [Epub ahead of print].

Müller, U., Steven Dick, A., Gela, K., Overton, W. F., and Zelazo, P. D. (2006). The role of negative priming in preschoolers' flexible rule use on the dimensional change card sort task. *Child Dev.* 77, 395–412. doi: 10.1111/j.1467-8624.2006.00878.x

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., and O'reilly, R. C. (2011). A unified framework for inhibitory control. *Trends Cogn. Sci.* 15, 453–459. doi: 10.1016/j.tics.2011.07.011

Munakata, Y., Snyder, H. R., and Chatham, C. H. (2012). Developing cognitive control three key transitions. *Curr. Dir. Psychol. Sci.* 21, 71–77. doi: 10.1177/0963721412436807

Perry, L. K., Samuelson, L. K., Malloy, L. M., and Schiffer, R. N. (2010). Learn locally, think globally exemplar variability supports higher-order generalization and word learning. *Psychol. Sci.* 21, 1894–1902. doi: 10.1177/0956797610389189

Reuss, H., Kiesel, A., Kunde, W., and Hommel, B. (2011). Unconscious activation of task sets. *Conscious. Cogn.* 20, 556–567. doi: 10.1016/j.concog.2011.02.014

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *J. Neurosci. Methods* 162, 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Richter, F. R., and Yeung, N. (2012). Memory and cognitive control in task switching. *Psychol. Sci.* 23, 1256–1263. doi: 10.1177/0956797612444613

Rogers, R. D., and Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *J. Exp. Psychol.* 124, 207. doi: 10.1037/0096-3445.124.2.207

St Clair-Thompson, H. L., and Gathercole, S. E. (2006). Executive functions and achievements in school: shifting, updating, inhibition, and working memory. *Q. J. Exp. Psychol.* 59, 745–759. doi: 10.1080/17470210500162854

Vandierendonck, A., Liefooghe, B., and Verbruggen, F. (2010). Task switching: interplay of reconfiguration and interference control. *Psychol. Bull.* 136, 601. doi: 10.1037/a0019791

Waszak, F., and Hommel, B. (2007). The costs and benefits of cross-task priming. *Mem. Cognit.* 35, 1175–1186. doi: 10.3758/BF03193487

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Blaye, A., and Chevalier, N. (2011). The role of goal representation in preschoolers' flexibility and inhibition. *J. Exp. Child Psychol.* 108, 469–483. doi: 10.1016/j.jecp.2010.09.006

Bull, R., Espy, K. A., and Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev. Neuropsychol.* 33, 205–228. doi: 10.1080/87565640801982312

Bull, R., and Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Dev. Neuropsychol.* 19, 273–293. doi: 10.1207/S15326942DN1903_3

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carroll, D. J., and Cragg, L. (2012). "Precursors to cognitive flexibility: rule-switching in two-year-olds," in *Paper presented at the Development of Executive Functions Workshop* (Utrecht, Netherlands).

Cepeda, N. J., Kramer, A. F., and Gonzalez De Sather, J. C. (2001). Changes in executive control across the life span: examination of task-switching performance. *Dev. Psychol.* 37, 715–730. doi: 10.1037/0012-1649.37.5.715

Chevalier, N., and Blaye, A. (2009). Setting goals to switch between tasks: effect of cue transparency on children's cognitive flexibility. *Dev. Psychol.* 45, 782–797. doi: 10.1037/a0015409

Chevalier, N., Blaye, A., Dufau, S., and Lucenet, J. (2010). What visual information do children and adults consider while switching between tasks? Eye-tracking investigation of cognitive flexibility development. *Dev. Psychol.* 46, 955–972. doi: 10.1037/a0019674

Cragg, L., and Chevalier, N. (2012). The processes underlying flexibility in childhood. *Q. J. Exp. Psychol.* 65, 209–232. doi: 10.1080/17470210903204618

Cragg, L., and Nation, K. (2009). Shifting development in mid-childhood: the influence of between-task interference. *Dev. Psychol.* 45, 1465–1479. doi: 10.1037/a0015360

Cragg, L., and Nation, K. (2010). Language and the development of cognitive control. *Top. Cogn. Sci.* 2, 631–642. doi: 10.1111/j.1756-8765.2009.01080.x

Crone, E. A., Donohue, S. E., Honomichl, R., Wendelken, C., and Bunge, S. A. (2006). Brain regions mediating flexible rule use during development. *J. Neurosci.* 26, 11239–11247. doi: 10.1523/JNEUROSCI.2165-06.2006

Dauvier, B., Chevalier, N., and Blaye, A. (2012). Using finite mixture of GLMs to explore variability in children's flexibility in a task-switching paradigm. *Cogn. Dev.* 27, 440–454. doi: 10.1016/j.cogdev.2012.07.004

Dibbets, P., and Jolles, J. (2006). The switch task for children: measuring mental flexibility in young children. *Cogn. Dev.* 21, 60–71. doi: 10.1016/j.cogdev.2005.09.004

Espy, K. A. (1997). The shape school: assessing executive function in preschool children. *Dev. Neuropsychol.* 13, 495–499. doi: 10.1080/87565649709540690

Fisher, A. V. (2011a). Automatic shifts of attention in the dimensional change card sort task: subtle changes in task materials lead to flexible switching. *J. Exp. Child Psychol.* 108, 211–219. doi: 10.1016/j.jecp.2010.07.001

Fisher, A. V. (2011b). Processing of perceptual information is more robust than processing of conceptual information in preschool-age children: evidence from costs of switching. *Cognition* 119, 253–264. doi: 10.1016/j.cognition.2011.01.015

Hommel, B., Kray, J., and Lindenberger, U. (2011). Feature integration across the lifespan: stickier stimulus–response bindings in children and older adults. *Front. Psychol.* 2:268. doi: 10.3389/fpsyg.2011.00268

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., and Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: age-related changes and individual differences. *Dev. Neuropsychol.* 28, 617–644. doi: 10.1207/s15326942dn2802_4

Hübner, M., Dreisbach, G., Haider, H., and Kluwe, R. H. (2003). Backward inhibition as a means of sequential task-set control: evidence for reduction of task competition. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 289–297. doi: 10.1037/0278-7393.29.2.289

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas Psychol.* 30, 190–200. doi: 10.1016/j.newideapsych.2011.11.001

Jacques, S., and Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): a measure of executive function in preschoolers. *Dev. Neuropsychol.* 20, 573–591. doi: 10.1207/S15326942DN2003_2

Karbach, J., and Kray, J. (2007). Developmental changes in switching between mental task sets: the influence of verbal labeling in childhood. *J. Cogn. Dev.* 8, 205–236. doi: 10.1080/15248370701202430

Kharitonova, M., Chien, S., Colunga, E., and Munakata, Y. (2009). More than a matter of getting 'unstuck': flexible thinkers use more abstract representations than perseverators. *Dev. Sci.* 12, 662–669. doi: 10.1111/j.1467-7687.2008.00799.x

Kharitonova, M., and Munakata, Y. (2011). The role of representations in executive function: investigating a developmental link between flexibility and abstraction. *Front. Psychol.* 2:347. doi: 10.3389/fpsyg.2011.00347

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., et al. (2010). Control and interference in task switching–a review. *Psychol. Bull.* 136, 849. doi: 10.1037/a0019842

Koch, I., and Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Mem. Cognit.* 34, 433–444. doi: 10.3758/BF03193420

Kray, J., and Eppinger, B. (2006). Effects of associative learning on age differences in task-set switching. *Acta Psychol.* 123, 187–203. doi: 10.1016/j.actpsy.2005.12.009

Kray, J., Karbach, J., and Blaye, A. (2012). The influence of stimulus-set size on developmental changes in cognitive control and conflict adaptation. *Acta Psychol.* 140, 119–128. doi: 10.1016/j.actpsy.2012.03.005

Mayr, U., and Bell, T. (2006). On how to be unpredictable evidence from the voluntary task-switching paradigm. *Psychol. Sci.* 17, 774–780. doi: 10.1111/j.1467-9280.2006.01781.x

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 22:1423. doi: 10.1037/0278-7393.22.6.1423

Monsell, S. (2003). Task switching. *Trends Cogn. Sci.* 7, 134–140. doi: 10.1016/S1364-6613(03)00028-7

Moriguchi, Y., and Hiraki, K. (2013). Behavioral and neural differences during two versions of cognitive shifting tasks in young children and adults. *Dev. Psychobiol.* doi: 10.1002/dev.21145/full. [Epub ahead of print].

Müller, U., Steven Dick, A., Gela, K., Overton, W. F., and Zelazo, P. D. (2006). The role of negative priming in preschoolers' flexible rule use on the dimensional change card sort task. *Child Dev.* 77, 395–412. doi: 10.1111/j.1467-8624.2006.00878.x

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., and O'reilly, R. C. (2011). A unified framework for inhibitory control. *Trends Cogn. Sci.* 15, 453–459. doi: 10.1016/j.tics.2011.07.011

Munakata, Y., Snyder, H. R., and Chatham, C. H. (2012). Developing cognitive control three key transitions. *Curr. Dir. Psychol. Sci.* 21, 71–77. doi: 10.1177/0963721412436807

Perry, L. K., Samuelson, L. K., Malloy, L. M., and Schiffer, R. N. (2010). Learn locally, think globally exemplar variability supports higher-order generalization and word learning. *Psychol. Sci.* 21, 1894–1902. doi: 10.1177/0956797610389189

Reuss, H., Kiesel, A., Kunde, W., and Hommel, B. (2011). Unconscious activation of task sets. *Conscious. Cogn.* 20, 556–567. doi: 10.1016/j.concog.2011.02.014

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *J. Neurosci. Methods* 162, 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Richter, F. R., and Yeung, N. (2012). Memory and cognitive control in task switching. *Psychol. Sci.* 23, 1256–1263. doi: 10.1177/0956797612444613

Rogers, R. D., and Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *J. Exp. Psychol.* 124, 207. doi: 10.1037/0096-3445.124.2.207

St Clair-Thompson, H. L., and Gathercole, S. E. (2006). Executive functions and achievements in school: shifting, updating, inhibition, and working memory. *Q. J. Exp. Psychol.* 59, 745–759. doi: 10.1080/17470210500162854

Vandierendonck, A., Liefooghe, B., and Verbruggen, F. (2010). Task switching: interplay of reconfiguration and interference control. *Psychol. Bull.* 136, 601. doi: 10.1037/a0019791

Waszak, F., and Hommel, B. (2007). The costs and benefits of cross-task priming. *Mem. Cognit.* 35, 1175–1186. doi: 10.3758/BF03193487

Waszak, F., Hommel, B., and Allport, A. (2003). Task-switching and long-term priming: role of episodic stimulus-task bindings in task-shift costs. *Cogn. Psychol.* 46, 361–413. doi: 10.1016/S0010-0285(02)00520-0

Waszak, F., Hommel, B., and Allport, A. (2004). Semantic generalization of stimulus-task bindings. *Psychon. Bull. Rev.* 11, 1027–1033. doi: 10.3758/BF03196732

Waszak, F., Hommel, B., and Allport, A. (2005). Interaction of task readiness and automatic retrieval in task switching: negative priming and competitor priming. *Mem. Cognit.* 33, 595–610. doi: 10.3758/BF03195327

Yeniad, N., Malda, M., Mesman, J., Van Ijzendoorn, M. H., and Pieper, S. (2013). Shifting ability predicts math and reading performance in children: a meta-analytical study. *Learn. Individ. Differ.* 23, 1–9. doi: 10.1016/j.lindif.2012.10.004

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Zelazo, P. D., Frye, D., and Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cogn. Dev.* 11, 37–63. doi: 10.1016/S0885-2014(96)90027-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 November 2013; accepted: 24 January 2014; published online: 12 February 2014.

Citation: FitzGibbon L, Cragg L and Carroll DJ (2014) Primed to be inflexible: the influence of set size on cognitive flexibility during childhood. Front. Psychol. ***5****:101. doi: 10.3389/fpsyg.2014.00101*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 FitzGibbon, Cragg and Carroll. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 09 September 2014
doi: 10.3389/fpsyg.2014.00968

"Trick or treat": the influence of incentives on developmental changes in feedback-based learning

Kerstin Unger[1]*, Berit Greulich[2] and Jutta Kray[2]

[1] *Developmental Cognitive Neuroscience Lab/Cognitive Neuroscience of Cognitive Control and Memory Lab, Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, USA*
[2] *Development of Language, Learning, and Action, Department of Psychology, Saarland University, Saarbrücken, Germany*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Marieke Esther Van Der Schaaf, Radboud University Nijmegen, Netherlands
Madeline Harms, University of Minnesota, USA

***Correspondence:**
Kerstin Unger, Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Box 1821, Providence, RI 02912, USA
e-mail: kerstin_unger@brown.edu

Developmental researchers have suggested that adolescents are characterized by stronger reward sensitivity than both children and younger adults. However, at this point, little is known about the extent to which developmental differences in incentive processing influence feedback-based learning. In this study, we applied an incentivized reinforcement learning task, in which errors resulted in losing money (loss condition), failure to gain money (gain condition), or neither (no-incentive condition). Children (10–11 years), younger adolescents (13–14 years), and older adolescents (15–17 years) performed this task while event-related potentials (ERPs) were recorded. We focused our analyses on two ERP correlates of error processing, the error negativity (Ne/ERN) and the error positivity (Pe) that are thought to reflect a rapid preconscious performance monitoring mechanism (Ne/ERN) and conscious detection and/or evaluation of response errors (Pe). Behaviorally, participants in all age groups responded more quickly and accurately to stimuli in gain and loss conditions than to those in the no-incentive condition. The performance data thus did not support the idea that incentives generally have a greater behavioral impact in adolescents than in children. While the Ne/ERN was not modulated by the incentive manipulation, both children and adolescents showed a larger Pe to errors in the gain condition compared to loss and no-incentive conditions. This is in contrast to results from adult studies, in which the Ne/ERN but not the Pe was enhanced for high-value errors, raising the possibility that motivational influences on performance monitoring might be reflected in the activity of separable neural systems in children and adolescents vs. adults. In contrast to the idea of higher reward/incentive sensitivity in adolescents, our findings suggest that incentives have similar effects on feedback-based learning from late childhood into late adolescence with no changes in preferences for "trick over treat."

Keywords: adolescence, childhood, incentives, performance monitoring, feedback-based learning, Ne/ERN, Pe

INTRODUCTION

Adolescence has often been characterized as a period of increased reward-seeking, risk-taking and impulsive behaviors (e.g., Casey et al., 2010; Somerville and Casey, 2010). A number of influential neurodevelopmental theories share the basic notion that this unique behavioral pattern reflects a relative imbalance in the maturation of the neural systems underlying (i) emotional and incentive-driven behavior, including subcortical structures such as the amygdala and the striatum, and (ii) cognitive and emotional control, including frontal regions such as the anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (Geier and Luna, 2009; Casey et al., 2010). Specifically, these models posit that the earlier maturation of subcortical systems can lead to a dominance of these structures over prefrontal control systems in guiding behavior, especially in situations involving salient motivational and/or social-affective cues. In contrast, prefrontal-subcortical interactions are more balanced in both children and young adults, due to a global immaturity (children) and maturity (adults) of the underlying neural circuitry. Accordingly, motivational cues like rewards are presumed to have a higher positive or negative impact on cognitive control in adolescence than at earlier or later stages of development.

In line with this view, numerous studies found that reward-related processing has a greater influence on decision-making in adolescents than in children or adults (e.g., Galvan et al., 2006; Cauffman et al., 2010; Somerville et al., 2010). While most of these studies focused on maladaptive effects of adolescents' hypersensitivity to incentives, it has recently been pointed out that pubertal changes in affective and social processing may also be associated with adaptive advantages. In particular, adolescents are thought to be biologically prepared to rapidly adjust to changing environmental conditions and hence should show a greater flexibility in the recruitment of cognitive control mechanisms to support motivational learning (cf. Crone and Dahl, 2012). Direct empirical tests of this proposal are scarce thus far. Using a reversal learning task, Van der Schaaf et al. (2011) demonstrated that adolescents are indeed better able to change their responses after unexpected rewarding and punishing outcomes

compared to both children and adults. Furthermore, a developmental neuroimaging study of reinforcement learning indicated that adolescents show exaggerated striatal responses to reward prediction errors, i.e., discrepancies between expected and actually obtained outcomes (Cohen et al., 2010). Although Cohen et al. (2010) did not find age differences in overall learning performance, adolescents responded more quickly to feedback stimuli indicating large reward compared to those associated with small reward.

In the present study, we sought to expand on previous research on interactions between motivational context and learning mechanisms across adolescence by examining the impact of appetitive and aversive motivational cues on error processing and error-related performance adjustments. We used the high temporal resolution of an event-related potentials (ERPs) to track earlier and later stages of error processing as reflected in two ERP correlates of performance monitoring, the error negativity (Ne; Falkenstein et al., 1990) or error-related negativity (ERN; Gehring et al., 1993) and the error positivity (Pe; Falkenstein et al., 1990).

The Ne is a fronto-centrally distributed negative deflection that peaks ∼30-100 ms after a participant's erroneous response. It is typically followed by the Pe, a slow positive wave that reaches its maximum between 200 and 400 ms after response-onset and exhibits a centroparietal scalp distribution (Falkenstein et al., 1990). While the Pe has been associated with deliberate, slower error evaluation processes, such as conscious error recognition and appraisal of the motivational significance of an error (Overbeek et al., 2005; Ridderinkhof et al., 2009; Steinhauser and Yeung, 2010), the Ne is thought to be a neural manifestation of a rapid internal response evaluation mechanism. More specifically, the Ne has been proposed to reflect the activity of a generic prefrontal performance monitoring system and to track learning-related changes in the evaluation and utilization of information about performance outcomes (Holroyd and Coles, 2002). Consistent with this notion, previous findings suggested a link between the Ne and error-induced behavioral adaptation during reinforcement learning (e.g., Frank et al., 2005; Gründler et al., 2009; Unger et al., 2012). Moreover, there is substantial evidence for motivational and affective influences on the Ne in adults (for a review, see Gehring et al., 2012). In particular, the Ne has been shown to be sensitive to the motivational value of an error (e.g., Gehring et al., 1993; Hajcak et al., 2005; Wiswede et al., 2009; Unger et al., 2012).

Developmental studies on error processing indicated that the Ne increases until mid to late adolescence (e.g., Davies et al., 2004; Ladouceur et al., 2007; for a recent review, see Ferdinand and Kray, 2014). Although some studies showed that a reliable Ne can be elicited in children as young as 5–7 years of age when using a simple Go-NoGo paradigm (e.g., Torpey et al., 2009), this component does not seem to develop until later ages for more complex tasks (e.g., Davies et al., 2004; Ladouceur et al., 2004). Eppinger et al. (2009) used a reinforcement learning paradigm to investigate age differences in error processing and found comparable accuracy rates and Ne amplitudes for children and adults in the easiest learning condition (valid feedback), while performance and Ne were increased in adults compared to children when the task was more difficult (invalid feedback). In addition, Eppinger et al. (2009) observed a larger Pe in children than adults, whereas other studies did not find age-related changes in this component (Davies et al., 2004; Ladouceur et al., 2004). The divergent findings may reflect differences in the paradigms used across studies (reinforcement learning vs. Eriksen Flanker task). Ladouceur et al. (2007), however, also reported an increase in Pe amplitude in late adolescents compared to adults, using a flanker task. This suggests that the neural processes involved in the generation of the Pe mature relatively early in development and amplitude modulations reflect age differences in error awareness or motivational significance of errors.

Available evidence regarding developmental changes of motivational influences on error processing as reflected in Ne and Pe is mixed. For instance, Kim et al. (2005) reported an increase in Ne when children performed a task while being observed by a peer compared to performing the task alone. Torpey et al. (2009), in contrast, failed to obtain significant differences in children's Ne and Pe amplitudes for high- vs. low-value errors. So far, only little is known on how motivational salience affects electrophysiological correlates of error processing and error-related performance adjustment across adolescence (cf. Ferdinand and Kray, 2014). Some insight has been gained from developmental research on the feedback-related negativity (FRN), an ERP component that has been hypothesized to reflect a rapid evaluation of the significance or value of external feedback stimuli and is thought to be functionally related to the Ne (Holroyd and Coles, 2002). Interestingly, the findings of these studies suggested that the neural system underlying the FRN differentiates less efficiently between good and bad events in adolescents compared to adults, for both small and large outcomes (e.g., Hämmerer et al., 2011; Zottoli and Grose-Fifer, 2012). The present investigation addressed the question of how motivational value affects internal rather than external indicators of performance errors. We applied an incentivized reinforcement learning paradigm in a sample of children and adolescents covering the age ranges of 10–11 years, 13–14 years, and 15–17 years. Incentive value was manipulated on a trial-to-trial basis in three different conditions: errors resulted in monetary loss (loss condition), failure to gain money (gain condition), or neither (no-incentive condition). On the basis of the theoretical considerations and previous findings outlined above, we expected adolescents to show better learning performance and larger Ne and/or Pe amplitudes in gain and loss conditions compared to the no-incentive condition, whereas incentive-related differences in behavioral and ERP measures should be less pronounced in children.

MATERIALS AND METHODS

PARTICIPANTS

A total of 64 children and adolescents participated in the study. Data from four participants (2 children, 2 adolescents) were excluded from analyses due to excessive artifacts in the EEG data. One child did not finish the session. The final sample thus included 59 participants from three age groups: 19 children (10–11 years, mean age = 11.02 years, 9 females), 20 mid-adolescents (13–14 years, mean age = 14.20 years, 10 females), and 20 late adolescents (15–17 years, mean age = 16.85 years, 10 females). The age ranges were chosen to (a) reflect

the development of performance monitoring and reinforcement learning from preadolescence to late adolescence (e.g., Galvan et al., 2006; Cauffman et al., 2010), (b) to cover the age period during which sensitivity to incentives has been shown to peak (e.g., Somerville and Casey, 2010), and (c) to be comparable to previous developmental studies using similar paradigms (e.g., Eppinger et al., 2009; Hämmerer et al., 2011). Participants were consented in accordance with the protocols approved by the local ethics committee of Saarland University and were paid 8 Euro per hour for participation. According to self-report, all had normal or corrected-to-normal vision, no history of neurological or psychiatric illness and did not use psychoactive medication or drugs. Five participants (1 child, 3 mid-adolescents, 1 late adolescent) were left-handed, all other participants were right-handed (Oldfield Questionnaire; Oldfield, 1971). The majority of children ($N = 16$), mid-adolescents ($N = 20$) and late adolescents ($N = 16$) were attending college-preparatory high school. The parents of children had an average 16.63 ($SD = 4.52$) years of education, the parents of younger and older adolescents had an average of 16.40 ($SD = 2.58$) and 15.39 ($SD = 2.87$) years of education, respectively.

STIMULI AND TASK

On each trial of the reinforcement learning task, participants saw a colored image of an object (Snodgrass and Vanderwart, 1980) and chose to press one of two response keys with the left and right index finger, respectively. Feedback was presented after each choice in the form of either a happy smiley (correct response) or a sad smiley (incorrect response). Stimuli were assigned to one of three incentive conditions (gain, loss, and no-incentive condition). Each imperative stimulus was preceded by a cue that indicated the incentive value of the upcoming target. The gain cue informed participants that they would win 37 euro cents if they responded correctly but 0 euro cents if they responded incorrectly or missed the response deadline (see Trial Procedure). Conversely, the loss cue indicated that participants would lose 0 euro cents if they responded correctly but 37 euro cents if the response was incorrect or too slow. On no-incentive trials, there was no chance to gain or lose money. The outcome of each trial was indicated by "+37," "+00," "−00," or "−37" signs shown together with the corresponding smiley on the feedback screen. At the end of the experiment, all participants received a performance-dependent monetary bonus (ranging between 5 and 10 Euros). In order to make the learning task more child-friendly, we constructed a cover story involving creatures living in a magic forest that have been transformed into different objects by a wizard. Participants were told that they have two magic wands (the two response keys) and should find out which one can be successfully used to free a given creature from the spell. They were further told that some creatures will reward successful retransformation with a monetary gift (gain condition), while others punish unsuccessful retransformations by taking away money (loss condition) or do neither (no-incentive condition).

TRIAL PROCEDURE

Each trial started with the incentive cue appearing in the center of the screen for 400 ms. After a 400 ms delay, a central fixation cross was displayed for 500 ms, followed by the presentation of the target stimulus for another 500 ms. Stimuli were presented on a light gray background. In order to minimize strategic adjustments in response speed across the incentive conditions, that is, more accurate but slower responding on gain and loss compared to no-incentive trials, we applied an adaptive response deadline. Depending on the proportion of time-out trials ($M = 0.04$, $SD = 0.01$), the response window was individually adjusted in steps of 100 ms within an overall range of 500–1500 ms (for a similar procedure, see Eppinger et al., 2009). After the key press, a blank screen was displayed for 500 ms and then visual feedback was provided for again 500 ms. If the participant failed to respond within the adaptive response time window, "Too Slow" feedback was shown. The next trial started after a randomly jittered 500 to 800 ms interval (see **Figure 1** for a schematic overview of the trial procedure).

EXPERIMENTAL PROCEDURE

The learning task consisted of a short practice block and 15 experimental blocks, with self-paced breaks every 30 trials. During the breaks, participants were presented with a feedback screen displaying the amount of money they had won so far (this value was always equal to or greater than zero, i.e., no negative scores were shown). Within one block, two stimuli were assigned to each incentive condition, yielding a total of six new stimuli per learning block. One of the two stimuli was mapped to the left response key and the other one to the right response key. Each stimulus was presented 10 times in pseudo-randomized order throughout the learning block, with the same stimulus appearing not more than two times in a row. The assignment of stimuli to incentive condition and response key was randomized across participants. The learning task took about 60 min to complete.

ELECTROPHYSIOLOGICAL RECORDING

The electroencephalogram (EEG) was recorded from 58 Ag/AgCl electrodes arranged according to the extended 10–20 system, referenced to the left mastoid, using Brain Amp DC Recorder (BrainVision recorder acquisition software). Data were sampled at 500 Hz in DC mode with a low-pass filter at 70 Hz. Impedances were kept below 5 kΩ. Electrodes placed on the outer canthi of the two eyes and on the infra- and supra-orbital ridges of the

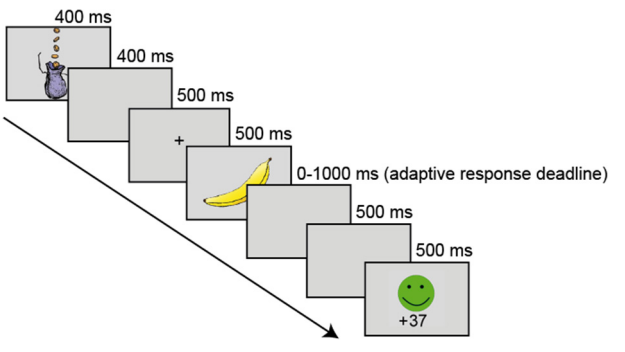


FIGURE 1 | Schematic overview of the trial procedure.

left eye recorded the horizontal and vertical electrooculograms. The data were re-referenced offline to the linked mastoids and band-pass filtered from 0.1 to 30 Hz. The impact of blinks and eye movements was corrected using an independent component analysis algorithm implemented in the BrainVision Analyzer Software Package (Brain products, Gilching, Germany). Trials containing EEG activity exceeding ±100 μV, changing more than 50 μV between samples or containing DC drifts were eliminated by a semiautomatic artifact inspection procedure.

DATA ANALYSES

Behavioral data analyses

Responses exceeding the adaptive deadline were excluded from further analyses. To examine the course of learning, each block was split into five bins. The bins were created according to the number of stimulus repetitions, i.e., Bin 1 contained first and second presentation of the respective stimuli, Bin 2 third and fourth presentation, and so on. Within each bin, mean reaction times (RTs) and accuracy rates were computed for the three incentive conditions. The number of time-out trials did not differ (a) between incentive conditions, (b) across bins, or (c) after correct vs. erroneous responses in either age group ($ps > 0.14$). On average, 58 trials were included in each bin per condition.

ERP analyses

Artifact-free EEG data were segmented relative to response onset and baseline-corrected using the average voltage in a −200 to −50 ms preresponse interval. We defined the Ne as mean amplitude in a 0–50 ms time window following the response. The interval was chosen to capture the peak of the Ne in each age group (see **Figure 4**). As in previous studies (e.g., Hajcak et al., 2004; Wiswede et al., 2009), the Pe was computed as the mean amplitude in a 200- to 400-ms postresponse interval. Both, Ne and Pe were scored at the three midline electrodes FCz, Cz, and Pz, separately for correct and incorrect trials. In order to make sure that ERPs for correct and erroneous responses included the same number of epochs, we randomly selected a subsample of correct trials based on the individual error trial counts in each condition. **Table 1** shows the mean number of EEG epochs that were used to quantify Ne and Pe in the three age groups.

Statistical analyses

Accuracy and ERP data were analyzed using repeated measures analyses of variance (ANOVAs). Whenever necessary, the Geisser-Greenhouse correction was applied (Geisser and Greenhouse, 1958) and corrected p-values are reported together with uncorrected degrees of freedom and the epsilon-values (ε). Planned comparisons were performed to decompose significant high-level interactions.

Table 1 | Mean number (standard deviation) of EEG epochs that were included in the calculation of Ne and Pe.

	Children	Mid-Adolescents	Late adolescents
Gain condition	69 (24)	45 (21)	57 (29)
Loss condition	70 (21)	45 (23)	60 (30)
No-incentive condition	74 (25)	50 (25)	68 (36)

RESULTS

BEHAVIORAL DATA

Reaction time and accuracy data were analyzed using ANOVAs with the between-subjects factor *Age group* (children, younger adolescents, older adolescents) and the within-subjects factors *Incentive condition* (gain, loss, and no-incentive) and *Bin* (Bins 1–5).

Reaction time

Figure 2 (see also **Table 2**) illustrates that RTs for all participants decreased with learning in each incentive condition [$F_{(4, 224)} = 6.35$, $p < 0.01$, $\varepsilon = 0.38$, $\eta_p^2 = 0.10$]. This was confirmed by a significant linear trend across bins [$F_{(1, 56)} = 8.03$, $p < 0.01$, $\eta_p^2 = 0.13$]. Moreover, RTs differed between the incentive conditions [$F_{(2, 112)} = 6.44$, $p < 0.01$, $\varepsilon = 0.70$, $\eta_p^2 = 0.10$] such that participants responded faster in gain and loss conditions compared to the no-incentive condition [$F_{(1, 56)} = 7.37$, $p < 0.01$, $\eta_p^2 = 0.12$]. As was indicated by an interaction of bin and incentive condition [$F_{(8, 448)} = 3.35$, $p < 0.01$, $\varepsilon = 0.77$, $\eta_p^2 = 0.06$], the learning-related speeding of responses varied for the three incentive conditions. Contrasts revealed that RTs decreased more rapidly in the gain compared to the loss condition [$F_{(1, 56)} = 17.92$, $p < 0.001$, $\eta_p^2 = 0.24$] as well as in gain and loss conditions compared to the no-incentive condition [$F_{(1, 56)} = 4.57$, $p < 0.05$, $\eta_p^2 = 0.08$]. There were also age differences in overall RT across age groups [$F_{(2, 56)} = 10.23$, $p < 0.001$, $\eta_p^2 = 0.27$]. Older adolescents responded faster than both younger adolescents and children [$F_{(1, 56)} = 19.80$, $p < 0.001$, $\eta_p^2 = 0.26$], whereas response latencies did not differ between the latter two age groups [$F < 1$, $p = 0.38$].

Accuracy

Accuracy learning curves for the three age groups in the three incentive conditions are shown in **Figure 3** and **Table 3**. Accuracy increased with age [$F_{(2, 56)} = 4.30$, $p < 0.05$, $\eta_p^2 = 0.13$] such that younger and older adolescents showed higher overall accuracy than children [$F_{(1, 56)} = 5.16$, $p < 0.05$, $\eta_p^2 = 0.08$] but did not significantly differ from each other ($p = 0.07$, $\eta_p^2 = 0.05$). As expected, all participants became more accurate across learning blocks [$F_{(4, 224)} = 302.389$, $p < 0.001$, $\varepsilon = 0.38$, $\eta_p^2 = 0.84$]. The course of learning, however, differed for the three age groups [$F_{(8, 224)} = 2.60$, $p < 0.05$, $\varepsilon = 0.59$, $\eta_p^2 = 0.09$]. Contrasts revealed stronger quadratic [$F_{(1, 56)} = 46.49$, $p < 0.001$, $\eta_p^2 = 0.45$] and cubic trends [$F_{(1, 56)} = 12.86$, $p < 0.01$, $\eta_p^2 = 0.18$] across bins in younger and older adolescents than in children, indicating that older participants learned faster and reached asymptote levels of accuracy earlier, while children's performance continued to improve more steadily throughout the learning blocks. Importantly, accuracy varied across the incentive condition [$F_{(2, 112)} = 10.56$, $p < 0.001$, $\varepsilon = 0.86$, $\eta_p^2 = 0.16$]. Participants showed better performance in gain and loss conditions compared to the no-incentive condition [$F_{(1, 56)} = 15.52$, $p < 0.001$, $\eta_p^2 = 0.22$], while accuracy scores did not differ reliably between the former two conditions ($F < 1$, $p = 0.55$). Similar to the RT data, however, there were no significant age differences in the influence of incentive-value on learning performance.

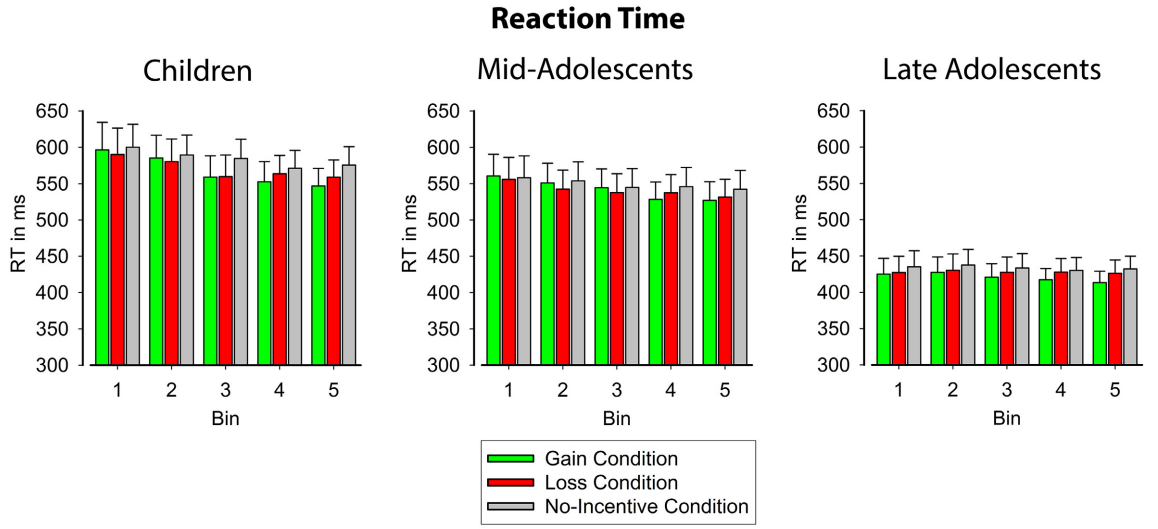


FIGURE 2 | Learning-related changes in reaction time for the three incentive conditions, displayed separately for children (left), younger adolescents (middle), and older adolescents (right). The x-axis shows the course of learning averaged into 5 bins. Error bars indicate standard error.

Table 2 | Mean reactions times (standard deviations) in ms for each condition and bin of the learning task.

Condition	Bin	Children	Mid-Adolescents	Late adolescents
Gain	1	597 (165)	561 (134)	425 (98)
	2	585 (137)	551 (122)	427 (95)
	3	559 (127)	544 (116)	421 (83)
	4	553 (120)	528 (106)	417 (68)
	5	547 (105)	527 (115)	413 (70)
Loss	1	590 (158)	546 (136)	427 (100)
	2	580 (135)	542 (118)	430 (100)
	3	560 (129)	538 (116)	427 (94)
	4	564 (110)	538 (111)	428 (84)
	5	559 (103)	532 (108)	426 (82)
No-incentive	1	600 (137)	558 (135)	435 (99)
	2	590 (119)	554 (117)	438 (97)
	3	585 (115)	545 (116)	434 (89)
	4	571 (106)	546 (119)	430 (80)
	5	576 (110)	542 (116)	432 (78)

ELECTROPHYSIOLOGICAL DATA

Figure 4 shows the ERPs to correct and erroneous responses at electrode site FCz, separately for the three incentive conditions for children, younger adolescents, and older adolescents. In all age groups, the Ne is evident as a fronto-centrally distributed negative deflection that is larger after erroneous than correct responses. The Ne increases with age, but is not clearly modulated by incentive condition. Following the Ne, the Pe can be observed as a positive deflection. In contrast to the Ne, the Pe seems to be smaller in older adolescents than in the two younger age groups and varies across the incentive conditions. Specifically, the Pe appears to be larger in the gain condition compared to loss and no-incentive condition.

Ne and Pe were subjected to separate ANOVAs involving the between-group factor *Age group* (children, younger adolescents, older adolescents) and the within-subjects factors *Incentive condition* (gain, loss, and no-incentive) and *Correctness* (correct vs. incorrect). In order to control for influences of RT differences between groups, we additionally ran ANVOVAs including mean response latencies as covariate. These analyses did not yield evidence that the ERP findings varied as a function of RT.

Error negativity

Analyses confirmed that the Ne was larger on incorrect compared to correct trials [$F_{(1,\ 56)} = 77.99$, $p < 0.001$, $\eta_p^2 = 0.58$] and increased from posterior to anterior sites [correctness × electrode: [$F_{(2,\ 112)} = 28.48, p < 0.001, \varepsilon = 0.70, \eta_p^2 = 0.34$]. As was indicated by a significant interaction of age group and correctness [$F_{(2,\ 56)} = 6.66, p < 0.01, \eta_p^2 = 0.19$], the Ne increased with age. Contrasts revealed that the amplitude difference between correct and erroneous trials was larger in the two adolescent groups compared to children [$F_{(1,\ 56)} = 5.73$, $p < 0.05$, $\eta_p^2 = 0.09$] as well as in 15–17-year-olds compared to 13–14-year olds [$F_{(1,\ 56)} = 12.52$, $p < 0.01$, $\eta_p^2 = 0.18$]. However, we found no evidence that the Ne reliably varied as a function of incentive condition in either age group ($ps > 0.27$).

Error positivity

As expected, there was a larger positivity on erroneous compared to correct trials [$F_{(1,\ 56)} = 39.08$, $p < 0.001$, $\eta_p^2 = 0.41$] and this amplitude difference was more pronounced at posterior than anterior scalp sites [correctness × electrode: $F_{(2,\ 112)} = 9.62$, $p < 0.01$, $\varepsilon = 0.67$, $\eta_p^2 = 0.18$]. Moreover, we found age-related differences in Pe amplitude [age group x correctness: $F_{(2,\ 56)} = 3.39$, $p < 0.05$, $\eta_p^2 = 0.11$]. Contrasts showed that the Pe was reduced in 15–17-year olds compared to the two younger age groups [$F_{(1,\ 56)} = 6.50$, $p < 0.05$, $\eta_p^2 = 0.10$], but did not significantly differ in 13–14-year-olds and 10–11-year-olds ($p = 0.63$). Most importantly, Pe amplitude differed between incentive conditions [$F_{(2,\ 112)} = 9.62, p < 0.001, \varepsilon = 0.89, \eta_p^2 = 0.15$], such that it was enhanced on gain trials compared to loss and no-incentive

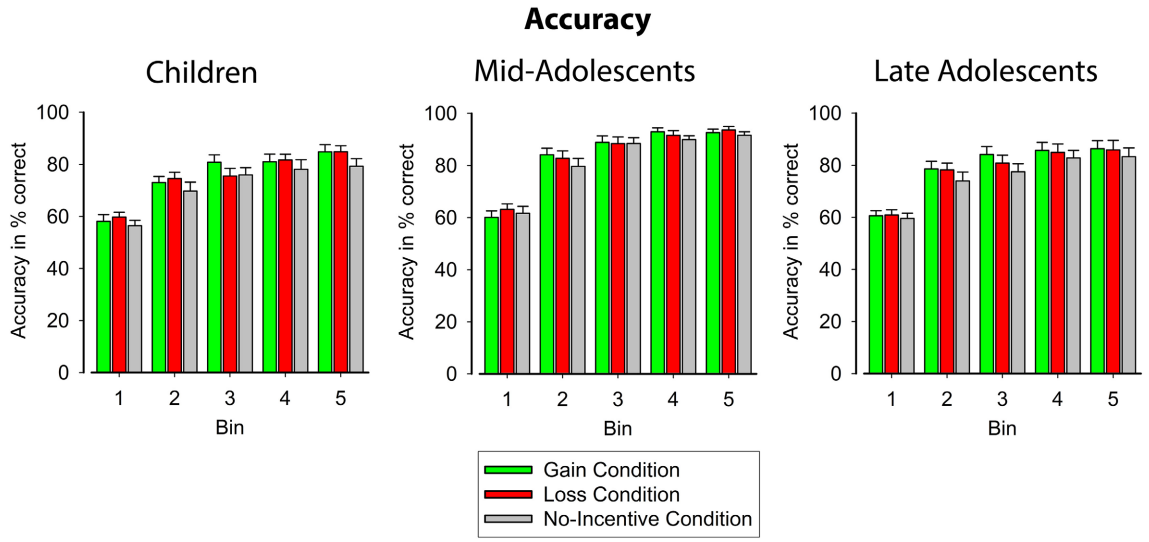


FIGURE 3 | Learning-related changes in accuracy rates for the three incentive conditions, displayed separately for children (left), younger adolescents (middle), and older adolescents (right). The x-axis shows the course of learning averaged into 5 bins. Error bars indicate standard error.

Table 3 | Mean accuracy rates (standard deviations) for each condition and bin of the learning task.

Condition	Bin	Children	Mid-adolescents	Late adolescents
Gain	1	0.58 (0.11)	0.60 (0.11)	0.61 (0.09)
	2	0.73 (0.10)	0.84 (0.11)	0.79 (0.13)
	3	0.81 (0.12)	0.89 (0.11)	0.84 (0.13)
	4	0.81 (0.13)	0.93 (0.07)	0.86 (0.14)
	5	0.84 (0.12)	0.93 (0.06)	0.86 (0.14)
Loss	1	0.60 (0.08)	0.63 (0.09)	0.61 (0.09)
	2	0.75 (0.11)	0.83 (0.13)	0.78 (0.12)
	3	0.76 (0.13)	0.89 (0.12)	0.81 (0.14)
	4	0.82 (0.10)	0.92 (0.08)	0.85 (0.14)
	5	0.85 (0.10)	0.94 (0.06)	0.86 (0.16)
No-incentive	1	0.56 (0.09)	0.62 (0.12)	0.60 (0.09)
	2	0.70 (0.15)	0.80 (0.14)	0.74 (0.15)
	3	0.76 (0.12)	0.88 (0.10)	0.78 (0.14)
	4	0.79 (0.16)	0.90 (0.08)	0.83 (0.13)
	5	0.79 (0.13)	0.92 (0.07)	0.83 (0.15)

trials [$F_{(1,\ 56)} = 9.98$, $p < 0.01$, $\eta_p^2 = 0.15$]. This effect, however, did not significantly vary as a function of age ($p = 0.38$).

DISCUSSION

In this study, we examined developmental differences in motivational influences on error processing—as reflected in Ne and Pe—and error-induced learning, comparing children (10–11 years), mid-adolescents (13–14 years) and late adolescents (15–17 years). We used an incentivized reinforcement learning task, in which errors resulted in losing money (loss condition), failure to gain money (gain condition), or neither (no-incentive condition).

Behaviorally, participants in all age groups responded more quickly and accurately to stimuli in gain and loss conditions than to those in the no-incentive condition. Thus, even 10–11-year-old children were able to efficiently use motivational cues in order to maximize outcomes of their task performance. The behavioral data, however, did not support the idea that motivational salience has a greater impact on learning-related performance adjustment in adolescents than in children. Instead, the current findings are in line with results from previous studies using "non-affective" decision-making paradigms that reported linear age-related performance improvements (e.g., Crone et al., 2008; Van Duijvenvoorde et al., 2008; Van den Bos et al., 2012). Similarly, Van der Schaaf et al. (2011) found that overall accuracy in a reversal-learning task (including reversal and non-reversal trials) increased with age, reaching asymptote at adolescence. However, the authors observed an inverted U-shaped relationship between age and performance on reversal trials, peaking at adolescence. Benefits of adolescents' aberrant sensitivity to salient motivational cues hence might be limited to situations that require a particularly high degree of flexibility, such as the need for rapid behavioral reversal in volatile environments. Thus, one reason for the failure to obtain non-linear age-related changes in the present study might have been that we used a deterministic learning task with fixed stimulus-response mappings, causing a predictable and stable environment.

In line with findings from previous developmental studies on performance monitoring (e.g., Davies et al., 2004; Santesso et al., 2006; Ladouceur et al., 2007), the Ne increased with age until late adolescence. The reduction of Ne amplitude in younger participants has been linked to the protracted structural and functional development of the medial prefrontal performance monitoring system (cf. Ladouceur et al., 2007; Torpey et al., 2009), especially the ACC (the putative neural source of the Ne; Gehring et al., 2012). Importantly, there is evidence to suggest that age difference in the Ne may be attributed to children's deficits in task performance rather than developmental changes in the neural

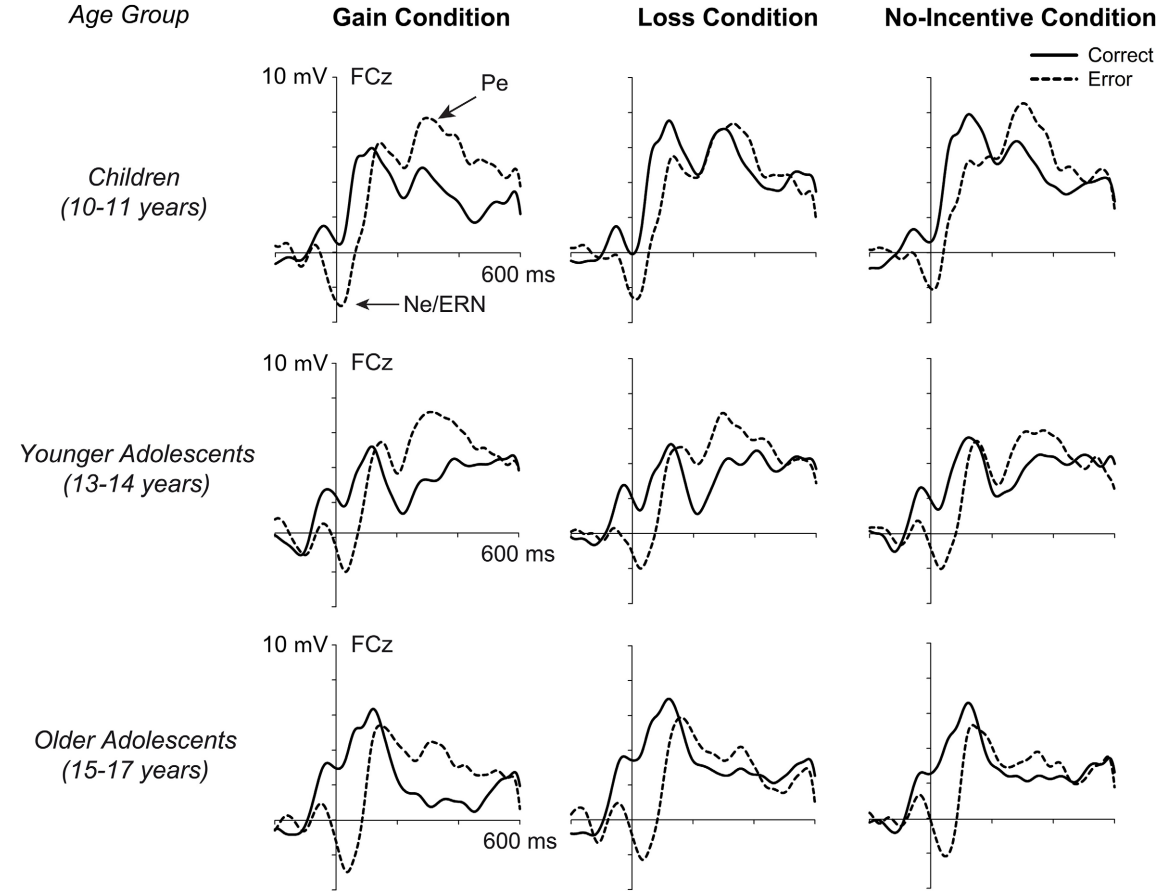


FIGURE 4 | Response-locked grand average ERPs for the three incentive conditions, displayed separately for correct responses (solid lines) and incorrect responses (dashed lines), for children (top), younger adolescents (middle), and older adolescents (bottom) at electrode FCz.

structures underlying performance monitoring (e.g., Eppinger et al., 2009). In the present study, however, we observed larger Ne amplitudes in older compared to younger adolescents in the absence of significant differences in overall accuracy. The current findings thus corroborate the view that neural systems underlying the Ne continue to develop throughout adolescence (Ladouceur et al., 2007).

This view would also be consistent with our observation that the Ne was not modulated by the incentive manipulation in children and adolescents, whereas previous studies demonstrated that the Ne is sensitive to such motivational influences in adults (e.g., Hajcak et al., 2005; Potts, 2011). Notably, we recently tested a sample of young adults using a highly similar reinforcement learning paradigm that included probabilistic instead of deterministic stimulus-response mappings and found a larger Ne in the loss condition compared to both gain and no-incentive conditions (Unger and Kray, unpublished data). The present results parallel findings by Torpey et al. (2009) in younger children (5–7 years), but contrast with the study by Kim et al. (2005), in which the presence of a peer during task performance was associated with an increase in the Ne in 7–8-year-olds. The latter finding raises the more general question of whether monetary incentives are sufficiently salient motivational cues for children and adolescents. Although the present data show incentive-related improvements in task performance, there is evidence to support the notion that social-affective incentives (e.g., peer admiration) may play a more prominent role in adolescence (Crone and Dahl, 2012). Despite the plausibility of this hypothesis, previous work demonstrated that monetary gains and losses have a substantial impact on decision making in adolescents (e.g., Van Duijvenvoorde et al., 2010).

While the Ne did not vary as a function of error-value, both children and adolescents showed a larger Pe to errors in the gain condition compared to loss and no-incentive conditions. This is in contrast to results from adult studies, in which the Ne but not the Pe was enhanced for high-value errors (e.g., Hajcak et al., 2005; Potts, 2011). Given that these two components are thought to reflect functionally dissociable mechanisms (Overbeek et al., 2005; Ridderinkhof et al., 2009; Steinhauser and Yeung, 2010), the present findings indicate that motivational influences on error processing qualitatively change across development. Interestingly, recent work established a specific link between error detection mechanisms and the Pe, whereas the Ne might be related to more general aspects of performance monitoring—such as conflict detection or tracking of error/reward likelihood—rather

than to error processing itself (Steinhauser and Yeung, 2010). Specifically, Steinhauser and Yeung (2010) suggested that the Pe reflects a process that feeds available evidence for the decision that a response error has occurred into an internal error detection system and hence may support deliberate performance adjustments. According to this view, the current findings indicate that participants were more certain about error commission, i.e., had stronger evidence that an error occurred—in the gain compared to loss and no-incentive conditions. While conscious error detection may have contributed to performance optimization on gain trials, it remains to be determined which mechanisms underlie improved performance in the loss condition. One possibility is that motivationally salient loss feedback is more robustly maintained in working memory (Frank et al., 2007).

Moreover, the inverse age-related changes in Pe (decrease) and Ne (increase) raise the interesting possibility that error-related remedial behaviors might rely on different mechanisms across development. While larger Pe amplitudes in children and mid-adolescents may reflect that younger participants gather more evidence to support conscious error detection (Steinhauser and Yeung, 2010), the enhanced Ne in older adolescents indicates stronger recruitment of more general performance monitoring mechanisms (e.g., conflict monitoring).

Some limitations of the current study should be noted. First, the learning paradigm might have been insensitive to unique features of motivational processing in adolescence. Clearly, future studies are needed to test whether subcortical mechanisms can exert beneficial influences on adolescents' performance in salient social-affective contexts or situations that require higher behavioral flexibility (e.g., volatile and uncertain environments). Second, one could argue that participants in the youngest age group were too close to adolescence and hence did not provide an appropriate "baseline." However, other studies that did find unique effects of motivational-affective variables on adolescents' decision making covered a similar age range (cf. Somerville et al., 2010). Moreover, sensitivity to incentives has been shown to peak between ages 14 and 16 (Somerville and Casey, 2010). Thus, it seems unlikely that limitations of age range account for the failure to obtain age differences in incentivized learning. Nonetheless, it is important to mention that age does not provide a precise measurement of pubertal development. Future studies thus should include a direct assessment of pubertal status.

Notably, comparisons of the present data with previous findings in adults suggest that (1) the medial prefrontal performance monitoring system underlying the Ne undergoes functional change until late adolescence and (2) incentive-related modulations in performance monitoring are reflected in the activity of at least partially dissociable neural systems in children and adolescents (modulations in the Pe) vs. young adults (modulations in the Ne) that may support more deliberate vs. automatic, preconscious forms of performance adjustment, respectively. Since our study did not include an adult group, these conclusions need to be tested by future studies applying the same learning paradigm in children, adolescents and young adults. Furthermore, future research may probe whether motivational factors influence the relationship between neural error signals and error-related performance adjustment on a single-trial basis.

To conclude, the present findings do not support the idea that incentives generally have a stronger impact on feedback-based learning in early and late adolescence than in late childhood. Instead, the behavioral data showed that both children and adolescents efficiently used incentive cues to optimize performance outcomes, with no systematic differences between salient reward (gain condition) and salient punishment (loss condition). However, the ERP data suggested that gain but not loss anticipation is associated with enhanced recruitment of error processing mechanisms as reflected in the Pe that are thought to support conscious error detection and deliberative performance adjustment.

ACKNOWLEDGMENTS

We thank Lisa Riedel for help with data collection and our participants and their families for making this research possible. This work was funded by the German Research Foundation under Grant DFG-IRTG-1457 and was conducted in the International Research Training Group "Adaptive Minds" hosted by Saarland University, Saarbrücken (Germany).

REFERENCES

Casey, B. J., Duhoux, S., and Cohen, M. M. (2010). Adolescence: what do transmission, transition, and translation have to do with it? *Neuron* 67, 749–760. doi: 10.1016/j.neuron.2010.08.033

Cauffman, E., Shulman, E. P., Steinberg, L., Claus, E., Banich, M. T., Graham, S., et al. (2010). Age differences in affective decision making as indexed by performance on the Iowa Gambling Task. *Dev. Psychol.* 46, 193–207. doi: 10.1037/a0016128

Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., et al. (2010). A unique adolescent response to reward prediction errors. *Nat. Neurosci.* 13, 669–671. doi: 10.1038/nn.2558

Crone, E. A., and Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nat. Rev. Neurosci.* 13, 636–650. doi: 10.1038/nrn3313

Crone, E. A., Zanolie, K., Van Leijenhorst, L., Westenberg, P. M., and Rombouts, S. A. (2008). Neural mechanisms supporting flexible performance adjustment during development. *Cogn. Affect. Behav. Neurosci.* 8, 165–177. doi: 10.3758/CABN.8.2.165

Davies, P. L., Segalowitz, S. J., and Gavin, W. J. (2004). Development of response-monitoring ERPs in 7- to 25-year-olds. *Dev. Neuropsychol.* 25, 355–376. doi: 10.1207/s15326942dn2503_6

Eppinger, B., Mock, B., and Kray, J. (2009). Developmental differences in learning and error processing: evidence from ERPs. *Psychophysiology* 46, 1043–1053. doi: 10.1111/j.1469-8986.2009.00838.x

Falkenstein, M., Hohnsbein, J., Hoorman, J., and Blanke, L. (1990). "Effects of errors in choice reaction tasks on the ERP under focused and divided attention," in *Psychophysiological Brain Research*, eds C. H. M. Brunia, A. W. K. Gaillard, and A. Kok (Tilburg: Tilburg University Press), 192–195.

Ferdinand, N. K., and Kray, J. (2014). Developmental changes in performance monitoring: how electrophysiological data can enhance our understanding of error and feedback processing in childhood and adolescence. *Behav. Brain Res.* 263, 122–132. doi: 10.1016/j.bbr.2014.01.029

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., and Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16311–16316. doi: 10.1073/pnas.0706111104

Frank, M. J., Woroch, B. S., and Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron* 47, 495–501. doi: 10.1016/j.neuron.2005.06.020

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., et al. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *J. Neurosci.* 26, 6885–6892. doi: 10.1523/JNEUROSCI.1062-06.2006

Gehring, J. W., Goss, B., Coles, M. G., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390. doi: 10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., Liu, Y., Orr, J. M., and Carp, J. (2012). "The error-related negativity (ERN/Ne)," in *Oxford Handbook of Event-Related Potential Components*, eds S. J. Luck and E. Kappenman (New York, NY: Oxford University Press), 231–291.

Geier, C., and Luna, B. (2009). The maturation of incentive processing and cognitive control. *Pharmacol. Biochem. Behav.* 93, 212–221. doi: 10.1016/j.pbb.2009.01.021

Geisser, S., and Greenhouse, S. W. (1958). An extension of box's results on the use of the F-distribution in multivariate analysis. *Ann. Math. Stat.* 29, 885–891. doi: 10.1214/aoms/1177706545

Gründler, T. O. J., Cavanagh, J. F., Figueroa, C. M., Frank, M. J., and Allen, J. J. B. (2009). Task-related dissociation in ERN amplitude as a function of obsessive-compulsive symptoms. *Neuropsychologia* 47, 1978–1987. doi: 10.1016/j.neuropsychologia.2009.03.010

Hajcak, G., McDonald, N., and Simons, R. F. (2004). Error-related psychophysiology andnegative affect. *Brain Cogn.* 56, 189-197. doi: 10.1016/j.bandc.2003.11.001

Hajcak, G., Moser, J. S., Yeung, N., and Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology* 42, 151-160. doi: 10.1111/j.1469-8986.2005.00270.x

Hämmerer, D., Li, S.-C., Müller, V., and Lindenberger, U. (2011). Life span differences in electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. *J. Cogn. Neurosci.* 23, 579–592. doi: 10.1162/jocn.2010.21475

Holroyd, C. B., and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709. doi: 10.1037/0033-295X.109.4.679

Kim, E. Y., Iwaki, N., Imashioya, H., Uno, H., and Fujita, T. (2005). Error-related negativity in children: effect of an observer. *Dev. Neuropsychol.* 28, 871–883. doi: 10.1207/s15326942dn2803_7

Ladouceur, C. D., Dahl, R. E., and Carter, C. S. (2004). ERP correlates of action monitoring in adolescence. *Ann. N. Y. Acad. Sci.* 1021, 329–336. doi: 10.1196/annals.1308.040

Ladouceur, C. D., Dahl, R. E., and Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: ERP and source localization. *Dev. Sci.* 10, 874–891. doi: 10.1111/j.1467-7687.2007.00639.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Overbeek, T. J. M., Nieuwenhuis, S., and Ridderinkhof, K. R. (2005). Dissociable components of error processing: on the functional significance of the Pe vis-a` -vis the ERN/Ne. *J. Psychophysiol.* 19, 319–329. doi: 10.1027/0269-8803.19.4.319

Potts, G. (2011). Impact of reward and punishment motivation on behavior monitoring as indexed by the error-related negativity. *Intl. J. Psychophysiol.* 81, 324–331. doi: 10.1016/j.ijpsycho.2011.07.020

Ridderinkhof, K. R., Ramautar, J. R., and Wijnen, J. G. (2009). To Pe or not to Pe: a P3-like ERP component reflecting the processing of response errors. *Psychophysiology* 46, 531–538. doi: 10.1111/j.1469-8986.2009.00790.x

Santesso, D. L., Segalowitz, S. J., and Schmidt, L. A. (2006). Error-related electrocortical responses in 10-year-old children and young adults. *Dev. Sci.* 9, 473–481. doi: 10.1111/j.1467-7687.2006.00514.x

Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol.* 6, 174–215.

Somerville, L. H., and Casey, B. (2010). Developmental neurobiology of cognitive control and motivational systems. *Curr. Opin. Neurobiol.* 20, 236–241. doi: 10.1016/j.conb.2010.01.006

Somerville, L. H., Jones, R. M., and Casey, B. J. (2010). A time of change: behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain Cogn.* 72, 124–133. doi: 10.1016/j.bandc.2009.07.003

Steinhauser, M., and Yeung, N. (2010). Decision processes in human performance monitoring. *J. Neurosci.* 30, 15643–15653. doi: 10.1523/JNEUROSCI.1899-10.2010

Torpey, D., Hajcak, G., and Klein, D. (2009). An examination of errorrelated brain activity and its modulation by error value in young children. *Dev. Neuropsychol.* 34, 749–761. doi: 10.1080/87565640903265103

Unger, K., Kray, J., and Mecklinger, A. (2012). Worse than feared? Failure induction modulates the electrophysiological signature of error monitoring during subsequent learning. *Cogn. Affect. Behav. Neurosci.* 12, 34–51. doi: 10.3758/s13415-011-0061-y

Van den Bos, W., Cohen, M. X., Kahnt, T., and Crone, E. A. (2012). Striatum–medial prefrontal cortex connectivity predicts developmental differences in reinforcement learning. *Cereb. Cortex* 22, 1247–1255. doi: 10.1093/cercor/bhr198

Van der Schaaf, M. E., Warmerdam, E., Crone, E. A., and Cools, R. (2011). Distinct linear and non-linear trajectories of reward and punishment reversal learning during development: relevance for dopamine's role in adolescent decision making. *Dev. Cogn. Neurosci.* 1, 578–590. doi: 10.1016/j.dcn.2011.06.007

Van Duijvenvoorde, A. C. K., Jansen, B. R. J., Visser, I., and Huizenga, H. M. (2010). Affective and cognitive decision-making in adolescents. *Dev. Neuropsychol.* 35, 539–554. doi: 10.1080/87565641.2010.494749

Van Duijvenvoorde, A. C., Zanolie, K., Rombouts, S. A., Raijmakers, M. E., and Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *J. Neurosci.* 28, 9495–9503. doi: 10.1523/JNEUROSCI.1485-08.2008

Wiswede, D., Münte, T. F., and Rüsseler, J. (2009). Negative affect induced by derogatory verbal feedback modulates the neural signature of error detection. *Soc. Cogn. Affect. Neurosci.* 4, 227–237. doi: 10.1093/scan/nsp015

Zottoli, T. M., and Grose-Fifer, J. (2012). The feedback-related negativity (FRN) in adolescents. *Psychophysiology* 49, 413–420. doi: 10.1111/j.1469-8986.2011.01312.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 February 2014; accepted: 13 August 2014; published online: 09 September 2014.

*Citation: Unger K, Greulich B and Kray J (2014) "Trick or treat": the influence of incentives on developmental changes in feedback-based learning. Front. Psychol. **5**:968. doi: 10.3389/fpsyg.2014.00968*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Unger, Greulich and Kray. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 24 December 2013
doi: 10.3389/fpsyg.2013.00959

Stuck in the moment: cognitive inflexibility in preschoolers following an extended time period

Carolina Garcia and Anthony Steven Dick*

Department of Psychology, Florida International University, Miami, FL, USA

Edited by:
Nicolas Chevalier, University of Colorado Boulder, USA

Reviewed by:
J. Bruce Morton, University of Western Ontario, Canada
Melody Wiseheart, York University, Canada

****Correspondence:***
Anthony Steven Dick, Department of Psychology, Florida International University, Modesto A. Maidique Campus AHC-4 454, 11200 S. W. 8th Street, Miami, FL 33199, USA
e-mail: adick@fiu.edu

Preschoolers display surprising inflexibility in problem solving, but seem to approach new challenges with a fresh slate. We provide evidence that while the former is true the latter is not. Here, we examined whether brief exposure to stimuli can influence children's problem solving following several weeks after first exposure to the stimuli. We administered a common executive function task, the Dimensional Change Card Sort, which requires children to sort picture cards by one dimension (e.g., color) and then switch to sort the same cards by a conflicting dimension (e.g., shape). After a week or after a month delay, we administered the second rule again. We found that 70% of preschoolers continued to sort by the initial sorting rule, even after a month delay, and even though they are explicitly told what to do. We discuss implications for theories of executive function development, and for classroom learning.

Keywords: cognitive flexibility, executive function, problem solving, Dimensional Change Card Sort, event binding

INTRODUCTION

Young children appear to pick up new knowledge with ease, but they display equally surprising inflexibility in problem solving situations. No better evidence for this exists than observing children's performance on a now-classic task, the Dimensional Change Card Sort (DCCS; Frye et al., 1995; Zelazo et al., 2003; Zelazo, 2006). In this task, children sort picture cards first by one dimension (e.g., color), and then by a second conflicting dimension (e.g., shape). Even though they are told exactly what to do (e.g., "Put the blue ones here and the yellow ones there"), when the rule changes (e.g., "Put the dogs here and the boats there"), younger preschoolers perseverate and sort by the initial rule. What remains unclear is whether this is a transient effect, or whether preschoolers who encounter conditions of conflict in a problem solving situation will carry with them the effects of that experience to the same problem solving situation if it occurs again in the more distant future.

Such an investigation must begin with an examination of why young children have difficulty with such an easy task. Several authors have focused on the processes that are called upon to overcome interference from the initial task set in the immediate context. For example, Frye et al. (1995) and Zelazo et al. (1996) proposed a Cognitive Complexity and Control (CCC) theory suggesting that in such task-switching situations children must organize the conflicting task rules into a hierarchical structure and apply that structure to determine which is the appropriate rule to use in the present context. This places a demand on working memory (Halford et al., 1998; Zelazo et al., 2003). Others suggest that, when the rule changes, children must inhibit attention to features of the object that were previously relevant in the immediate past, but are now irrelevant after the rule switch (Kirkham et al., 2003; Diamond and Kirkham, 2005; Diamond et al., 2005). This is proposed to require mature inhibitory control over the immediate context.

These theories, however, remain agnostic about how processes of working memory and inhibition interface with long term memory of prior experiences with the objects under study. In the laboratory setting this is less relevant because the stimuli used to study task-switching are often novel for the children. However, in everyday problem solving situations encountered in school and home learning settings, because it is cost-prohibitive to regularly replace the teaching tools, preschoolers often encounter objects repeatedly over the course of the school year. For example, preschool children might learn to categorize blocks or toys by their shape when they are learning shape categories, and they might then move on to learn to categorize those same objects by color. The theoretical accounts just described suggest that, in order to shift to categorize the objects by color, children would have to recruit additional cognitive resources (whether working memory or inhibitory control) to consider the same objects under a different category. However, these theories make no predictions about whether the children would have to recruit these additional resources if they encountered the objects several weeks or months after the initial experience with the objects.

Work with adults suggests that they would, and that (1) tasks carry their history with them and when task stimuli are faced again there is a re-establishment of the previous task set; (2) stimuli acquire associations with the tasks in which they occur; (3) facing the same stimuli in different tasks produces cognitive costs; (4) these effects can be detected even after long intervening time periods (Waszak et al., 2003). Additional theoretical and empirical research on memory and priming in children also suggests that children in such settings would encounter what amounts to a task-switching situation, even if the initial encounter with the

objects occurred many weeks before. For example, from a theoretical perspective such a prediction would likely fit within an active-latent connectionist account of children's performance on the DCCS. This model draws a distinction between different memory systems that are presumably engaged in tasks like the DCCS: a slow, "latent" memory system implemented in the form of connections between processing units, established over time during the pre-switch phase of the DCCS, and a fast "active" memory system implemented in the form of units capable of self-sustaining activity (Munakata, 2001; Morton and Munakata, 2002; Blackwell et al., 2009). The continued interference of the pre-switch dimension would be predicted under this model if the latent memory traces persist for long periods of time (Yerys and Munakata, 2006).

The prediction could also be made under the revised CCC theory (CCC-r; Zelazo et al., 2003; Müller et al., 2006), which specifies that negative priming of the irrelevant dimension during the pre-switch phase contributes to difficulty in the post-switch phase. Negative priming describes the disruption or slowing of a response to a stimulus that has previously been ignored (Fox, 1995; May et al., 1995). Applied to the DCCS, the theory suggests that during the pre-switch phase there is suppression or inhibition of the competing distractor (negative priming) such that when the distractor stimuli become relevant in the post-switch, the previous suppression must be overcome by application of a higher-order rule. Early theories of negative priming suggested that inhibition of the competing distractor was a transient effect (Tipper, 1985; Neill and Westberry, 1987). However, priming effects of seemingly innocuous stimuli are known to occur after months (Sloman et al., 1988; Maylor, 1998) and even years (Mitchell, 2006), even in preschool children (Drummey and Newcombe, 1995), and subsequent research has shown that negative priming effects can also last a considerable amount of time (Tipper, 2001 for review). For example, DeSchepper and Treisman (1996) observed negative priming in a selective attention task in adults after 1 month between the presentation of the prime and the subsequent presentation of the probe. This suggests that the memory traces formed even within a single experience with an object can last at full strength across several weeks of temporal delay. If negative priming contributes to difficulty on the DCCS, it would be expected that the initial experience with the stimulus properties would still contribute to interference weeks later.

There is evidence that negative priming affects performance on the DCCS over short delays of 10 min (Müller et al., 2006), but these findings provide only initial evidence that children continue to "swim against the current" if they previously encountered a problem and failed to solve it, even if that encounter occurred in the more distant past. If the effects are found to be very long lasting, it would have implications for understanding mechanisms underlying children's cognitive inflexibility, and it would also have implications for how parents and teachers structure problem solving situations in education settings. That is, if the same materials are used to teach conflicting concepts, such a finding would suggest that some concept learning situations can actually require additional cognitive resources to overcome interference from the prior processing episode. Further, this might be the case even if the initial experience with the event was brief, and occurred several weeks or months in the past.

To investigate this issue more thoroughly, we examined whether even brief exposure to stimuli can influence problem solving following a significant intervening time after the first exposure. We administered a second post-switch phase to the DCCS following either a week or a month delay. Based on the priming literature we have reviewed, we predicted that the initial experience would have long-term effects on children's cognitive flexibility.

MATERIALS AND METHODS

PARTICIPANTS

We tested sixty-two 3–5-year-old children ($M = 4.0\ y$; $\mathrm{SD} = 0.58\ y$) from Miami-Dade, FL, USA preschools. Half participated in one of two conditions (1 week and 1 month), and did not differ on age across conditions, $t(60) = -0.215$, $p = 0.83$. Children were bilingual and were tested in their dominant language (assessed by parent questionnaire and verified by pretest; nine children were verified only by pretest). Bilingualism has been related to improved cognitive flexibility on the DCCS (Bialystok, 1999). We restricted the study to bilingual children because the majority of children in Miami-Dade preschools are bilingual, and we wanted to maintain homogeneity of the sample on this factor. Testing took place in the preschools. To control for possible effects of context, testing took place in the same location for all phases of the task.

GENERAL DESIGN

The design was a between subjects design with Delay (1 week and 1 month) as a single factor with two levels.

We followed Zelazo (2006), with the exception that we added a second post-switch phase following the first post-switch, after either 1 week or 1 month. We used two target cards (e.g., blue dog and yellow boat) and 10 test cards (e.g., yellow dogs and blue boats). Children were randomly assigned to sort by either color or shape first, and this was not associated with task success, Fisher's exact $p = 0.10$.

We attached one target card to each of the trays. We explained the rules for the pre-switch phase (e.g., "All the yellow ones go here, and all the blue ones go there.") and the child watched as the first practice trial was sorted. We then asked the child to sort a card, and we provided feedback to make sure they understood the instructions. We proceeded to the pre-switch phase and asked the children to sort the remaining eight cards and place them face down into the tray. On each trial, children were reminded of the rules and asked "Here's a (e.g., yellow one), where does this go in the (e.g., color) game?" After eight trials, we administered the post-switch rules – e.g., "Now we are going to play a new game. We are not going to play the color game anymore. We are going to play the shape game. In the shape game, all the dogs go here, and all the boats go there." Children then sorted eight post-switch trials.

We returned to the school after an intervening period of either 1 week or 1 month. Here we only presented one set of rules identical to those given in the post-switch phase in the initial encounter. For each trial, we repeated the rule, but no feedback was given. We administered knowledge questions after the post-switch phase (Zelazo et al., 1996).

Table 1 | Pass–fail rates for the first and second post-switch phases for each condition on the Dimensional Change Card Sort task.

		1 Week Delay	
		Post-switch Time 2	
		Pass	Fail
Post-switch Time 1	Pass	19	0
	Fail	3	9
	Total	22	9
		1 Month Delay	
		Post-switch Time 2	
		Pass	**Fail**
Post-switch Time 1	Pass	14	0
	Fail	6	11
	Total	20	11

RESULTS

Passing was defined as sorting four or more (out of eight cards) correctly. Consistent with prior work on the DCCS (Zelazo et al., 2003), the majority of children (90%) sorted correctly either all the cards, or none of the cards. All children included in analysis passed the pre-switch phase and answered the knowledge questions correctly. We first established that the groups did not differ at the first visit. Of the 31 children in each group, 12 failed the first post-switch for the 1-week delay condition, and 17 failed the first post-switch for the 1-month delay condition. These failure rates did not differ across condition, Fisher's exact $p = 0.37$, which indicates that the pre-post differences we report after the delay cannot be attributed to differences between groups during the first visit.

Of primary interest was whether the delay would help responding. This was tested separately for both conditions, with the null hypothesis that the passing rates for the first and second post-switch phases would be equal. The pass–fail rates for each condition are given in **Table 1**. For the week condition, 9 of 12 (75%) of children who failed the first post-switch also failed the second, McNemar $\chi^2(31) = 1.33$, $p = 0.25$, suggesting no significant difference in passing rate after the week delay. Specifically, only 25% of children benefited from the delay of 1 week. For the month condition, fewer children failed. Specifically, 11 of 17 (65%) failed both post-switch phases, McNemar $\chi^2(31) = 4.17$, $p = 0.04$, which shows a significant benefit for the month delay. Thus, 35% of children passed the second post-switch after failing the first post-switch. However, the result must be interpreted within the context that only a minority of children in either condition benefited from the delay. In fact, a large percentage of children (70% across conditions) who failed the first post-switch also failed the second after a considerable intervening time period.

DISCUSSION

We administered the DCCS with a second post-switch following either a week or a month delay. We showed that while a long delay of 1 month was sufficient to facilitate shifting to the novel dimension on the DCCS, even with this considerable intervening time period, 65% of children still failed to pass the task. While it is remarkable to observe young children's difficulty with the standard DCCS, our results extend this well-known finding to demonstrate that even brief exposure to simple stimuli can have a marked effect on children's success in simple problem solving situations many weeks later. As we discuss, these findings have implications for theories of developing cognitive flexibility (Garon et al., 2008; Cragg and Chevalier, 2012) and for problem solving situations in educational settings.

Theoretically, two models proposed to explain performance on the DCCS can be affected by our findings because they propose processes that could potentially change, in terms of their influence on performance, over the delay period. The first are models that emphasize the role of priming and negative priming in developing cognitive flexibility (Allport and Wylie, 2000; Zelazo et al., 2003; Müller et al., 2006; Chevalier and Blaye, 2008). In particular, there is growing evidence that negative priming contributes to children's difficulty on the DCCS (Zelazo et al., 2003; Müller et al., 2006) and on similar set-shifting tasks (Chevalier and Blaye, 2008; Dick, 2012). Further, these negative priming effects can be detected after only one or two conflicting stimulus presentations, and persist over a 10-min intervening time period (Müller et al., 2006; Experiment 4). Work with adults suggests that negative priming is not transient and can actually persist over a considerable time period (DeSchepper and Treisman, 1996), but until now this has not been shown in children. Our results connect well with the adult data in this respect, showing that negative priming of the irrelevant values may still influence the problem solving situation after several weeks (Tipper, 2001). However, the possibility that our results reflect the effects of prior negative priming would have to be confirmed, as it is possible that the findings could also be explained by interference from the previously relevant stimulus values (Kirkham et al., 2003; Chevalier and Blaye, 2008; Dick, 2012).

A second model of DCCS performance, the "active-latent" model, can also incorporate the findings we present here. As we reviewed in the Introduction, applied to the DCCS, the active-latent model proposes that flexible behavior is understood in terms of the relative strengths of "active" and "latent" memory traces (Munakata, 2001; Morton and Munakata, 2002; Blackwell et al., 2009). These memory traces are proposed to be graded in terms of the strength of the representation, and further are established over time during the pre-switch phase of the DCCS. Applied to this model, our data suggest that the latent memory trace is quite resilient in the face of considerable exposure to the stimulus values (e.g., yellow and blue) in other settings. In other words, encountering the stimulus values in another setting does not appear to "break the bond" between those particular values and the values with which they are associated during the pre-switch phase of the DCCS (e.g., boats and dogs).

The findings also fit well with models proposed for adults in task-switching situations, which suggest that tasks carry their history, and can elicit switch costs when the stimuli from the previous task are encountered in a different task (Waszak et al., 2003). The robust influence, even after long delays, of prior experience with the specific stimulus may be attributed to the retrieval

of "event-bindings" comprised of the object-task-action feature associations of the initial experience (Allport and Wylie, 2000; Zmigrod and Hommel, 2013). Research shows that even repeating parts of a previous feature combination can lead to the retrieval of all components of that combination (Kahneman et al., 1992; Hommel, 2004 for review). For example, functional magnetic resonance imaging (fMRI) studies have shown that repeating a particular stimulus feature reactivates areas of the brain involved in both the representation of that feature and the representation of features that co-occurred with that feature. Thus, Keizer et al. (2008) showed that, if presented after a subject perceives a face moving across the screen in a particular direction, seeing a house move in that same direction will activate areas of cortex sensitive to features of the house and the face, even though the face is not immediately present. Kühn et al. (2011) further showed that this binding affect applies to the response as well. That is, they showed that repeating a stimulus feature leads to the neural activation of regions involved in the response, and reactivation of regions involved in a different stimulus feature that accompanied the response. Some evidence suggests that combining features (e.g., shape and location) in memory is less efficient in children (Lorsbach and Reimer, 2005; Cowan et al., 2006), but our data suggest that, even if this is the case, the binding is adequate enough to affect problem solving after a long intervening time period.

Event- or feature-binding during problem solving facilitates rapid responding to stimuli that are experienced again in the future, but this is beneficial only if the task remains the same. If the task changes, especially if it conflicts with the previous task and uses the same objects, it will lead to interference and reduced facilitation. One can readily see how this would be relevant to educational settings. If this binding is as robust as our data imply, using the same objects to emphasize a particular concept or stimulus feature would be beneficial if the concept or stimulus feature is the same, but would potentially impede learning if the concept or stimulus feature is in conflict with the previous learning episode. For example, if a teacher or parent is trying to teach a preschooler the names for shapes, it might impede learning if the same objects were previously used to teach about colors because the previous event-bindings would be recalled. Further, this could occur long after the instructor would expect the child to remember the previous experience with the objects. It remains to be determined how important this is in actual educational settings–this requires additional research. However, at a minimum our results should help educators understand the challenges that preschoolers face beyond those that are apparent in the immediate situation.

If these implications are valid, one can ask what steps can be taken to minimize the interference effects of the prior task. One option is corrective feedback, which is shown to have a significant influence on maintenance of correct responses in testing situations (Bangert-Drowns et al., 1991 for review). This is particularly important in situations in which students make mistakes, as persistence of incorrect responding is known to increase the acquisition of false knowledge (Roediger and Marsh, 2005; Butler et al., 2006). In these situations, feedback promotes the learning of correct responses (Butler et al., 2008) and predicts better performance on subsequent tests (McDaniel and Fisher, 1991; Butler and Roediger, 2008).

The effects of feedback have been specifically assessed on tasks assessing cognitive flexibility such as the DCCS. One form of feedback is labeling of the dimensions, and a number of studies have shown positive effects of labeling the relevant properties on children's performance on the DCCS (Towse et al., 2000; Kirkham et al., 2003; Yerys and Munakata, 2006). However, both the timing of the feedback, and the nature of the labeling, affects responding (Yerys and Munakata, 2006), and some researchers have failed to find a beneficial effect of labeling on the DCCS (Müller et al., 2008). Age-related change in response to feedback is also indicated in tasks assessing cognitive flexibility. In one study, Chevalier et al. (2009) modeled the effects of feedback on an inductive task similar to the DCCS. They showed that children's responses are affected differently by different kinds of feedback. For example, early in the task children responded well to positive feedback for the relevant color, but not negative feedback for the irrelevant colors. However, this effect changes as the task proceeds through various phases of dimensional shifts–that is, the response to feedback changes across phases of the task. Chevalier et al. (2009) also showed that age modulated feedback processing efficiency as children progressed through the task. Such findings indicate that feedback provided in situations that require cognitive flexibility can have different effects depending on the complexity of the task, how the feedback relates to the prior experiences with the stimuli, and the age of the child.

In summary, the study we report revealed a surprising finding – for preschoolers, even very brief exposure to conflicting stimuli can influence the response to those stimuli if the problem solving situation is encountered again after a long intervening time period. Evidence for the resilience of the initial representation of the stimuli should be incorporated into existing theoretical models of cognitive flexibility. Further, the results should inform future work on how to structure learning in educational settings where the available resources often require teaching sometimes conflicting concepts using the same stimuli. Our data suggest that even waiting a long time between learning opportunities is insufficient to "wash out" prior experience with the task stimuli.

AUTHOR CONTRIBUTIONS

Anthony Steven Dick developed the study concept and design. Carolina Garcia collected the data. Anthony Steven Dick analyzed the data. Carolina Garcia and Anthony Steven Dick wrote the paper. Both authors approved the final version of the paper for submission.

REFERENCES

Allport, A., and Wylie, G. (2000). "Task switching, stimulus-response bindings, and negative priming," in *Control of Cognitive Processes: Attention and Performance XVIII,* eds S. Monsell and J. Driver (Cambridge, MA: MIT Press), 35–70.

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., and Morgan, M. (1991). The instructional effect of feedback in test-like events. *Rev. Educ. Res.* 61, 213–238. doi: 10.3102/00346543061002213

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* 70, 636–644. doi: 10.1111/1467-8624.00046

Blackwell, K. A., Cepeda, N. J., and Munakata, Y. (2009). When simple things are meaningful: working memory strength predicts children's cognitive flexibility. *J. Exp. Child Psychol.* 103, 241–249. doi:10.1016/j.jecp.2009.01.002

Butler, A. C., and Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem. Cognit.* 36, 604–616. doi: 10.3758/MC.36.3.604

Butler, A. C., Karpicke, J. D., and Roediger, H. L. (2008). Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 918–928. doi:10.1037/0278-7393.34.4.918

Butler, A. C., Marsh, E. J., Goode, M. K., and Roediger, H. L. (2006). When additional multiple-choice lures aid versus hinder later memory. *Appl. Cogn. Psychol.* 20, 941–956. doi: 10.1002/acp.1239

Chevalier, N., and Blaye, A. (2008). Cognitive flexibility in preschoolers: the role of representation activation and maintenance. *Dev. Sci.* 11, 339–353. doi: 10.1111/j.1467-7687.2008.00679.x

Chevalier, N., Dauvier, B., and Blaye, A. (2009). Preschoolers' use of feedback for flexible behavior: insights from a computational model. *J. Exp. Child Psychol.* 103, 251–267. doi:10.1016/j.jecp.2009.03.002

Cowan, N., Naveh-Benjamin, M., Kilb, A., and Saults, J. S. (2006). Life-span development of visual working memory: when is feature binding difficult? *Dev. Psychol.* 42, 1089–1102. doi: 10.1037/0012-1649.42.6.1089

Cragg, L., and Chevalier, N. (2012). The processes underlying flexibility in childhood. *Q. J. Exp. Psychol.* 209–232. doi: 10.1080/17470210903204618

DeSchepper, B., and Treisman, A. (1996). Visual memory for novel shapes: implicit coding without attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 27–47. doi: 10.1037/0278-7393.22.1.27

Diamond, A., Carlson, S. M., and Beck, D. M. (2005). Preschool children's performance in task switching on the dimensional change card sort task: separating the dimensions aids the ability to switch. *Dev. Neuropsychol.* 28, 689–729. doi:10.1207/s15326942dn2802_7

Diamond, A., and Kirkham, N. (2005). Not quite as grown-up as we like to think: parallels between cognition in childhood and adulthood. *Psychol. Sci.* 16, 291–297. doi:10.1111/j.0956-7976.2005.01530.x

Dick, A. S. (2012). Sources of cognitive inflexibility in set-shifting tasks: insights into developmental theories from adult data. *J.Cogn. Dev.* 13, 82–110. doi: 10.1080/15248372.2011.573516

Drummey, A. B., and Newcombe, N. (1995). Remembering versus knowing the past: children's explicit and implicit memories for pictures. *J. Exp. Child Psychol.* 59, 549–565. doi: 10.1006/jecp.1995.1025

Fox, E. (1995). Negative priming from ignored distractors in visual selection: a review. *Psychon. Bull. Rev.* 2, 145–173. doi:10.3758/BF03210958

Frye, D., Zelazo, P. D., and Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cogn. Dev.* 10, 483–527. doi: 10.1016/0885-2014(95)90024-1

Garon, N., Bryson, S. E., and Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi:10.1037/0033-2909.134.1.31

Halford, G. S., Wilson, W. H., and Phillips, S. (1998). Processing capacity defined by relational complexity: implications for comparative, developmental, and cognitive psychology. *Behav. Brain Sci.* 21, 803–831.[Discussion]. doi: 10.1017/S0140525X98001769

Hommel, B. (2004). Event files: feature binding in and across perception and action. *Trends Cogn. Sci.* 8, 494–500. doi: 10.1016/j.tics.2004.08.007

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24, 175–219. doi: 10.1016/0010-0285(92)90007-O

Keizer, A. W., Nieuwenhuis, S., Colzato, L. S., Teeuwisse, W., Rombouts, S. A., and Hommel, B. (2008). When moving faces activate the house area: an fFRI study of object-file retrieval. *Behav. Brain Funct.* 4, 50. doi:10.1186/1744-9081-4-50

Kirkham, N. Z., Cruess, L., and Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Dev. Sci.* 6, 449–476. doi: 10.1111/1467-7687.00300

Kühn, S., Keizer, A. W., Colzato, L. S., Rombouts, S. A., and Hommel, B. (2011). The neural underpinnings of event-file management: evidence for stimulus-induced activation of and competition among stimulus-response bindings. *J. Cogn. Neurosci.* 23, 896–904. doi:10.1162/jocn.2010.21485

Lorsbach, T. C., and Reimer, J. F. (2005). Feature binding in children and young adults. *J. Genet. Psychol.* 166, 313–328. doi: 10.3200/GNTP.166.3.313-328

Maylor, E. A. (1998). Retrieving names in old age: short- and (very) long-term effects of repetition. *Mem. Cognit.* 26, 309–319. doi: 10.3758/BF03201142

May, C. P., Kane, M. J., and Hasher, L. (1995). Determinants of negative priming. *Psychol. Bull.* 118, 35–54. doi: 10.1037/0033-2909.118.1.35

McDaniel, M. A., and Fisher, R. (1991). Tests and test feedback as learning sources. *Contemp. Educ. Psychol.* 16, 192–201. doi: 10.1016/0361-476X(91)90037-L

Mitchell, D. B. (2006). Nonconscious priming after 17 years: invulnerable implicit memory? *Psychol. Sci.* 17, 925–929. doi: 10.1111/j.1467-9280.2006.01805.x

Morton, J. B., and Munakata, Y. (2002). Active versus latent representations: a neural network model of perseveration, dissociation, and decalage. *Dev. Psychobiol.* 40, 255–265. doi: 10.1002/dev.10033

Müller, U., Dick, A. S., Gela, K., Overton, W. F., and Zelazo, P. D. (2006). The role of negative priming in preschoolers' flexible rule use on the dimensional change card sort task. *Child Dev.* 77, 395–412. doi: 10.1111/j.1467-8624.2006.00878.x

Müller, U., Zelazo, P. D., Lurye, L. E., and Liebermann, D. P. (2008). The effect of labeling on preschool children's performance in the dimensional change card sort task. *Cogn. Dev.* 23, 395–408. doi: 10.1016/j.cogdev.2008.06.001

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends Cogn. Sci.* 5, 309–315. doi: 10.1016/S1364-6613(00)01682-X

Neill, W. T., and Westberry, R. L. (1987). Selective attention and the suppression of cognitive noise. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 327–334. doi: 10.1037/0278-7393.13.2.327

Roediger, H. L., and Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1155–1159. doi:10.1037/0278-7393.31.5.1155

Sloman, S. A., Hayman, C. A. G., Ohta, N., Law, J., and Tulving, E. (1988). Forgetting in primed fragment completion. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 223–239. doi: 10.1037/0278-7393.14.2.223

Tipper, S. P. (1985). The negative priming effect: inhibitory priming by ignored objects. *Q. J. Exp. Psychol. A Hum. Exp. Psychol.* 37, 571–590. doi: 10.1080/14640748508400920

Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *Q. J. Exp. Psychol. A Hum. Exp. Psychol.* 54, 321–343. doi:10.1080/713755969

Towse, J. N., Redbond, J., Houston-Price, C. M. T., and Cook, S. (2000). Understanding the dimensional change card sort: perspectives from task success and failure. *Cogn. Dev.* 15, 347–365. doi: 10.1016/S0885-2014(00)00021-6

Waszak, F., Hommel, B., and Allport, A. (2003). Task-switching and long-term priming: role of episodic stimulus-task bindings in task-shift costs. *Cognit. Psychol.* 46, 361–413. doi: 10.1016/S0010-0285(02)00520-0

Yerys, B. E., and Munakata, Y. (2006). When labels hurt but novelty helps: children's perseveration and flexibility in a card-sorting task. *Child Dev.* 77, 1589–1607. doi: 10.1111/j.1467-8624.2006.00961.x

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi:10.1038/nprot.2006.46

Zelazo, P. D., Frye, D., and Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cogn. Dev.* 11, 37–63. doi: 10.1016/S0885-2014(96)90027-1

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., et al. (2003). The development of executive function in early childhood. *Monogr. Soc. Res. Child Dev.* 68, vii–137. doi: 10.1111/j.0037-976X.2003.00269.x

Zmigrod, S., and Hommel, B. (2013). Feature integration across multimodal perception and action: a review. *Multisens. Res.* 26, 143–157. doi: 10.1163/22134808-00002390

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 September 2013; paper pending published: 26 September 2013; accepted: 05 December 2013; published online: 24 December 2013.
*Citation: Garcia C and Dick AS (2013) Stuck in the moment: cognitive inflexibility in preschoolers following an extended time period. Front. Psychol. **4**:959. doi: 10.3389/fpsyg.2013.00959*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2013 Garcia and Dick. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 14 March 2014
doi: 10.3389/fpsyg.2014.00213

An effect of inhibitory load in children while keeping working memory load constant

Andy Wright and Adele Diamond*

Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

Edited by:
Philip D. Zelazo, University of Minnesota, USA

Reviewed by:
Jennifer Van Reet, Providence College, USA
Ruth Ford, Anglia Ruskin University, UK

****Correspondence:***
Adele Diamond, Department of Psychiatry, University of British Columbia, 2255 Wesbrook Mall, Vancouver, BC V6T 2A1, Canada
e-mail: adele.diamond@ubc.ca

Children are slower and more error-prone when the correct response is counter to their initial inclination (incongruent trials) than when they just need to do what comes naturally (congruent trials). Children are almost always tested on a congruent-trial block and then on an incongruent-trial block. That order of testing makes it impossible to determine whether worse performance on incongruent trials is due to the need to inhibit a pre-potent response, the need to clear the rule for Block 1 from working memory, some other demand of task-switching, or some combination of these. However, if the congruent block and incongruent blocks each have only one rule (e.g., "press on the same side as the stimulus" for congruent trials and "press on the side opposite the stimulus" for incongruent trials, as on the hearts and flowers task) and children's performance when the incongruent block is presented *first* is fully comparable to their performance when it is presented second, the only possible explanation for their worse performance on incongruent versus congruent trials would seem to be the added inhibitory demand on incongruent trials. Certainly, worse performance on Block 1 would not be due to inefficient clearing of working memory or task-switching demands. We tested 96 children (49 girls) 6–10 years of age on the hearts and flowers test with order of congruent and incongruent blocks counterbalanced across children. Children were slower and made more errors on incongruent trials regardless of task order. We expected task-switching demands to account for some of the variance, but to our surprise, performance was fully comparable on the incongruent block whether it came first or second. These results indicate that increasing inhibitory demands alone is sufficient to impair children's performance in the face of no change in working memory demands, suggesting that inhibition is a separate mental function from working memory.

Keywords: executive function, inhibitory control, self-regulation, cognitive control, executive control, spatial Stroop task, Simon task, stimulus-response compatibility

INTRODUCTION

It is hotly debated whether working memory and inhibitory controls are separable or not. Many argue that working memory is all that is required; no need to posit a separate inhibitory control ability (Cohen et al., 2002; Egner and Hirsch, 2005; Nieuwenhuis and Yeung, 2005; Hanania and Smith, 2010; Munakata et al., 2011; Chatham et al., 2012). Others posit that inhibitory control is an ability in its own right, separate from working memory (e.g., Gernsbacher, 1993; Levy and Anderson, 2002; Gazzaley et al., 2005; Leroux et al., 2006; Diamond, 2009; Zanto and Gazzaley, 2009).

When performing tasks that require working memory and inhibitory control, children are slower and make more errors on incongruent (incompatible) blocks than on congruent (compatible) ones. Each block may have only one rule but incongruent blocks add an inhibitory demand. When the incongruent block follows a congruent one, poorer performance on the incongruent block could easily be due to problems in efficiently clearing the congruent rule from working memory. Thus the working memory demand might be greater on Block 2 than on Block 1. However, when the incongruent block is presented first, worse performance on the incongruent block compared to the congruent one should be attributable to the greater inhibitory demand in the incongruent block. Such performance, if found, would seem to provide evidence in favor of working memory and inhibitory control being separable. To our knowledge, the study reported here is the first to present the incongruent-trial block before the congruent one to children.

For this study we wanted a task (a) that requires working memory (not just memory maintenance or short-term memory), (b) where the congruent and incongruent blocks each present only one rule to hold and manipulate in working memory, and (c) where there is clear empirical evidence that incongruent trials require a response counter to subjects' first inclination or response tendency [i.e., that "response inhibition," a component of "inhibitory control" (Diamond, 2013) is required]. The hearts and flowers task fit that bill.

The hearts and flowers task (previously called the dots task; Davidson et al., 2006; Diamond et al., 2007) is a hybrid combining elements of Simon and spatial Stroop tasks. For congruent trials, subjects are to obey the rule, "Press on the same side as the stimulus." For incongruent trials, subjects are to follow the rule, "Press on the side opposite the stimulus." Both of those blocks require working memory because we do not have "same side" or "opposite

side" hands (we have right and left hands); on each trial those rules must be translated into which hand to use (requiring that subjects mentally work with the rule they are holding in mind). This is an important difference between Simon tasks and the hearts and flowers task. Simon tasks require short-term memory, but not working memory, because they require simply holding two rules in mind ("For Stimulus A, press on the right" and "For Stimulus B, press on the left"), not mentally manipulating that information in any way.

Short-term memory involves only "memory maintenance," only holding information in mind (as required by a forward digit span task where you need to repeat back information you just heard in the order in which you heard it). Working memory, in contrast, requires memory maintenance plus working with the information you are holding in mind (as would be required if you need to repeat back information you just heard re-ordering it according to size, numerical or alphabetical order, or some other criterion; Baddeley, 1992; Petrides, 1994, 1995; D'Esposito et al., 1995, 1998; Owen et al., 1996; Cohen et al., 1997; Smith and Jonides, 1999; Smith et al., 1998).

Children at all ages that were tested (4–13 years) and young adults perform significantly better (fewer errors and faster responding) on the Simon task (with the memory demand of only holding information in mind) than on the hearts and flowers task [with the memory demands of holding information in mind plus manipulating that information (translating "same side" and "opposite side" into "right hand" or "left hand")]; see **Figure 1**.

People have a pre-potent tendency to respond toward a stimulus (Fitts and Seeger, 1953; Simon and Rudell, 1967; Lu and Proctor, 1995; Kornblum et al., 1999; Hommel et al., 2004; Hommel, 2011). That must be inhibited when the stimulus and its associated response are on opposite sides (incongruent trials). Adults and children are slower and make more errors when the stimulus appears on the side opposite its associated response than when stimulus appears on the same side as its associated response (called the Simon effect, the spatial incompatibility effect, or stimulus-response incompatibility; *adults*: Lu and Proctor, 1995; Kornblum et al., 1999; Kunde and Stocker, 2002; Hommel et al., 2004; Hommel, 2011; *children*: Gerardi-Coulton, 2000; Davidson et al., 2006; Mullane et al., 2009). Indeed, when monkeys are to respond away from a visual stimulus, the neuronal population vector in primary motor cortex (coding the direction of planned movement) initially points toward the stimulus and only then shifts to the required direction (showing a pre-potent tendency at the neuronal level to respond toward a stimulus; to do otherwise requires that that impulse be inhibited; Georgopoulos et al., 1989; Georgopoulos, 1994). This has been seen in humans using lateralized motor-readiness evoked potentials (Valle-Inclan, 1996) and event-related optical imaging (EROS; DeSoto et al., 2001). DeSoto et al. (2001) showed that incongruent trials elicit simultaneous activation of both motor cortices (necessitating the need for one to be inhibited) whereas congruent trials elicit brain activity in only the motor cortex associated with the response.

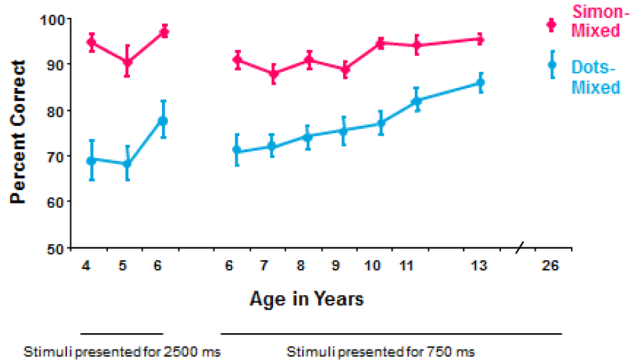


FIGURE 1 | Comparison of performance on the mixed conditions of the hearts and flowers task with dot stimuli and of a Simon task. This is based on within-subject comparisons of 314 participants (roughly 30 per age; equal numbers of males and females) tested on both tasks using the same equipment and same timing parameters (Davidson et al., 2006). At every age, participants were significantly faster and significantly more accurate on the Simon task. The dot stimuli were a gray disk and a black-and-white-striped disk; that is the only difference between the older dots version of the task and the current hearts and flowers task. The rules for the Simon task were, "If you see a butterfly, press the button on the left, whether the butterfly appears on the left or right. If you see a frog, press the button on to the right, whether the frog appears on the left or right."

Thus, the hearts and flowers task met all three of our criteria. In the standard hearts and flowers task, participants are instructed (a) to press the response button on the same side (left or right) as the stimulus (a red heart) on Block 1 (the congruent block), (b) to press the response button on the side opposite the stimulus (a red flower) on Block 2 (the incongruent block), and (c) to flexibly switch between those two rules on Block 3 where the stimulus might be a heart or flower (the mixed block). Participants of every age that has been tested (4–13 years, plus young adults) are slower and make more errors on the mixed block (Davidson et al., 2006). Young adults, however, are as fast and accurate on the incongruent block as they are on the congruent one. In contrast, children of all ages tested (4–13 years) are slower and make more errors on the incongruent block than the congruent one (Davidson et al., 2006).

The hearts and flowers task has been used to demonstrate executive function gains from the *Tools of the Mind* preschool curriculum (Diamond et al., 2007), to provide the first demonstration in children of a difference in executive function performance by COMT genotype (Diamond et al., 2004), and to demonstrate a sex difference in which version of the COMT gene is more beneficial for executive functions (Evans et al., 2009). It has been shown to accurately assess executive functions in both typically developing children and children with Down syndrome (Edgin et al., 2010). Zaitchik et al. (2013) found that the hearts and flowers task, but not several other tasks in their executive function battery, predicted their composite measure of vitalist biology as it is constructed by children (as predicted) controlling for age and IQ. The relation between hearts and flowers performance and on-the-face-of-it task-demands on their biology measures also held up (e.g., inhibitory control as indexed by hearts and flowers predicted animism judgments more strongly than purely factual knowledge about bodily function).

Using the hearts and flowers task, the present study tested two competing hypotheses:

(1) Children might err on the incongruent block because of the addition of an inhibitory demand – the need to resist responding on the same side as the stimulus, responding on the opposite side instead. For the congruent block children need only do what comes naturally, but for the incongruent block they must inhibit that and do the opposite. Thus, Hypothesis 1 is that children make more errors and take more time to respond on the incongruent block because of their immature ability to exercise inhibitory control.
(2) Perhaps, however, it is the task-switching requirement (and need to efficiently delete the rule for Block 1 from working memory when performing Block 2) that gives children difficulty. The incongruent block routinely follows the congruent one on most tasks, including the hearts and flowers task. Hypothesis 2 is that it is the difficulty of switching from the rule to always press on the same side as the stimulus to the rule to always pressing on the side opposite the stimulus that accounts for children's slower response times and increased errors on the incongruent block. We know that switching from one rule to another can be difficult even for adults, and especially for children (Hartman and Hasher, 1991; Allport and Wylie, 2000; Monsell and Driver, 2000; Cepeda et al., 2001; Zelazo et al., 2003; Crone et al., 2006; Yeung et al., 2006).

It may be that children do not wipe their mental slate clean when they begin Block 2, and so are still holding the now-irrelevant rule from Block 1 in mind. That would mean that the memory load for them on Block 2 would be greater because they would be holding in mind both the congruent and incongruent rules. If that is the case, then reversing the order in which the congruent and incongruent blocks are presented should get rid of poorer performance on the incongruent block. Hypothesis 1, on the other hand, leads to the prediction that reversing the order would do nothing to diminish the gap in children's performance on Blocks 1 and 2 (they would still be slower and less accurate on the incongruent block, even if it came first, because the inhibitory-control demand would be the same).

In a between-subjects design we tested half the children at each age with the congruent block first and half with the incongruent block first on the hearts and flowers task.

MATERIALS AND METHODS

PARTICIPANTS

Data were obtained from 96 children, ranging in age from 6 to 10 years (49% male, 51% female; see **Table 1**), from public elementary schools throughout the Lower Mainland of BC, Canada. Participants were recruited through their schools and 95% were tested at their school. The other five children were tested at our child development lab at the University of British Columbia.

The majority of participants who provided ethnic information were Caucasian of European descent (52%), 16% were of East Asian descent (most were Chinese), 12% were of South Asian descent (most were Indian), and the rest were of other ethnic backgrounds. All were fluent in English. Informed consent was obtained from a parents of each child, and informed assent was obtained from each child, before testing. All participants received a small present for their participation.

PROCEDURE

Within each age × gender grouping, half the participants were randomly assigned to get the congruent block first and half to get the incongruent block first. Participants were tested individually in a quiet room while wearing noise cancelation headphones. The stimuli were presented on a Dell 43 cm touchscreen monitor attached to an IBM ThinkPad Lenovo T6 laptop computer. The hearts and flowers task was administered using Presentation® software.

Participants held a handlebar with both hands to keep the distance from their hands to the response buttons constant. They were instructed to use only their pointer finger to press the response button on the screen (see **Figure 2**).

All participants completed a button practice task before moving onto hearts and flowers. Two response buttons appeared on the touchscreen monitor for the practice task. Children were to press a response button as soon as they saw a smiley face appear on it. This task provided baseline choice-reaction time data as well as serving to acclimate children to using the handlebars and to pressing the left and right response buttons on screen. Children were corrected if they reached across the midline to respond. They were also corrected if they left their finger on the monitor after their response, did not keep their hands on the handlebars before the smiley face appeared, or did not replace their finger on the handlebars after pushing the button.

Table 1 | Number of participants within each age and gender group.

Age group (years)	Mean age (years)	SD	*N*	Gender		Location of testing	
				Female (%)	Male (%)	Our lab	School
6	6.50	0.31	18	50	50	2	16
7	7.64	0.24	15	47	53	0	15
8	8.50	0.34	16	56	44	1	15
9	9.49	0.29	23	43	57	2	21
10	10.35	0.23	24	54	46	0	24
Totals			96	51	49	5	91

FIGURE 2 | A child performing the hearts and flowers test using a touchscreen monitor and handlebars.

The same procedure for the hearts and flowers task was used as previously reported (Davidson et al., 2006; Diamond et al., 2007). On each trial, a red heart or a red flower appeared on either the left or right side of the screen. A correct response to the heart was to press the response box on the touchscreen monitor on the same side as the heart. A correct response to the flower was to press the response box on the side opposite the flower.

On each trial, a horizontal rectangle (6 cm × 18 cm) was presented in the center of the screen. An orienting crosshair was presented for 500 ms at center fixation at the outset of each trial, and then disappeared, replaced 500 ms later by a stimulus on the left or right. One stimulus was presented per trial. The stimulus was presented for 750 ms to children ≥7 years of age and for 1500 ms to children 6 years of age. [These timing parameters had been determined to be age appropriate by Davidson et al. (2006)]. Each test block was preceded by instructions and a demonstration of the task followed by a practice block. Understanding of the rule was demonstrated by getting at least three of the four trials correct in the practice block. If understanding was not demonstrated on the first practice block, the child was instructed again and given another practice block (two children in the incongruent-first condition and two in the congruent-first condition needed a second practice block). No participant in the study failed to pass practice. The congruent and incongruent test blocks consisted of 12 trials each. There were 33 trials in the mixed block. Trials in each block were presented in the same pseudo-random order to each child.

RESULTS

The two dependent measures were speed [reaction time (RT)] and accuracy [percentage of correct responses]. Trials with RTs faster than 250 ms were excluded for being too fast to have been in response to the stimulus (resulted in 5 trials being excluded). RTs 2 standard deviations above or below a subject's mean were also excluded from analyses for being outliers (3 trials excluded). Percentage of correct responses was calculated by dividing the number of correct responses by the total number of responses (excluding the exceptions just mentioned). Only correct trials were used in calculating a child's mean RT in each test block.

RESULTS FOR SPEED OF RESPONDING

No significant difference was found between RTs during button practice (our baseline measure of choice RT) of children who received the incongruent block first and children who received the congruent block first [ANCOVA: $F(1,89) = 0.83$, ns] controlling for age, gender, and ethnicity. That is, there was no difference in baseline speed between children who received one order of presentation or the other. Choice RT did not vary by gender [$F(1,89) = 0.053$, ns] or ethnicity [$F(3,89) = 0.21$, ns]. Older children of course had faster choice RTs than younger children (all subjects: $F(1,78) = 12.34$, $p < 0.001$; only those receiving the same timing parameters [excludes 6-year-olds (who were given longer to respond)]: $F(1,71) = 4.17$, $p < 0.05$).

Excluding the 6-year-olds, who were given much more time to respond, RTs declined significantly over age for only the mixed condition of the hearts and flowers task [mixed block: $F(3,74) = 2.639$, $p < 0.05$; congruent block: $F(3,74) = 0.69$, ns; incongruent block: $F(3,74) = 0.05$, ns]. Since age is a continuous variable, we also examined this using multiple regression (controlling for gender and ethnicity) and received comparable results [mixed block: $F(1,72) = 5.07$, $p < 0.03$; congruent block: $F(1,72) = 0.70$, ns; incongruent block: $F(1,76) = 0.02$, ns]. RTs did not differ for any block by gender or ethnicity.

To compare the differences in RT between the congruent block and incongruent block for each order of testing, an ANOVA was conducted with order in which the congruent and incongruent blocks were presented as between-subject factors and block type (congruent, incongruent) as a within-subject factor; age, gender, and ethnicity were not included given the absence of any significant effects for those variables or their interactions. For the congruent-first condition (the order usually used for the hearts and flowers task) RTs in the congruent block (Mean = 596.44 ms, SD = 116.81) were significantly faster than in the incongruent block (Mean = 725.17 ms, SD = 167.17): $F(1,45) = 25.79$, $p < 0.001$. For the incongruent-first condition, RTs in the congruent block (Mean = 600.79 ms, SD = 138.05) were also significantly faster than in the incongruent block (Mean = 721.37 ms, SD = 180.89): $F(1,47) = 19.02$, $p < 0.001$. In both orders of testing, *at every age*, children responded faster in the congruent block than in the incongruent one (see **Figure 3A**). The difference between RTs on incongruent and congruent blocks did not differ by the order in which the blocks were presented: $F(1,93) = 1.41$, ns. *At no age* did the within-child difference in speed on the two blocks differ significantly by order of presentation (see **Figure 3B**). All of the above also held for each gender and for each ethnic group analyzed separately.

An insignificant *p*-value is not always sufficient for concluding that two conditions are equivalent (Lesaffre, 2008). Equivalence between the congruent-first and incongruent-first conditions on both congruent and incongruent trials was tested by setting a 95% confidence interval around the mean RT for each block in the congruent-first condition, and specifying equivalence as the RTs in the incongruent-first condition being within plus or

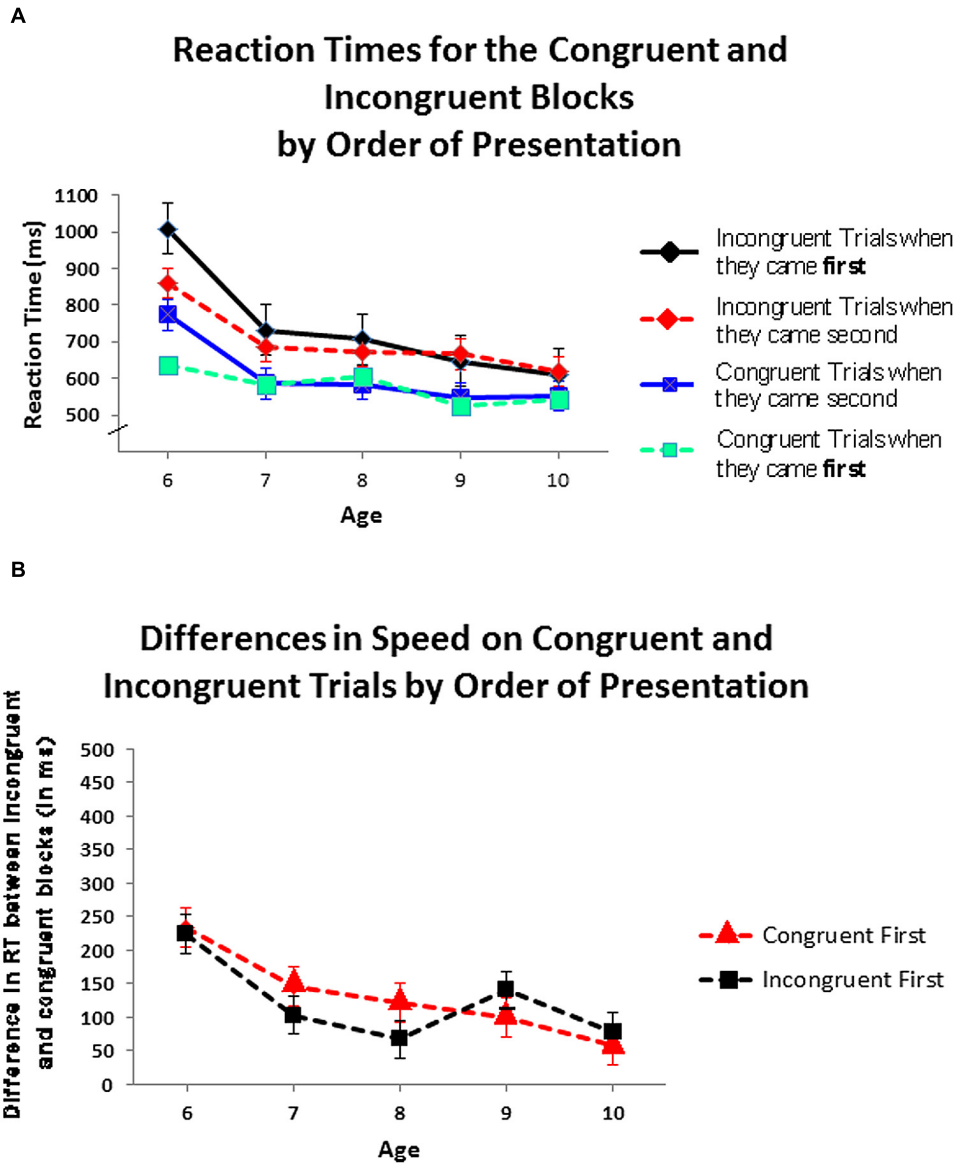


FIGURE 3 | (A) Speed of responding on the congruent and incongruent blocks by order of presentation. **(B)** Differences in speed on incongruent and congruent trial blocks by order of presentation.Values greater than zero, as all are, indicate that children responded faster on congruent than on incongruent blocks.

minus 1%. The mean RTs on the congruent block (**Figure 4A**) and incongruent block (**Figure 4B**) in the incongruent-first condition fell within the specified interval of equivalence when compared with the mean RT on the corresponding blocks in the congruent-first condition. This means that the mean RTs were equivalent for congruent trials whether they came first or second and the mean RTs were also equivalent for incongruent trials regardless of the order in which they we presented. The distribution of RTs was also similar. The equivalence of the difference in RT between the congruent and incongruent blocks in both congruent-first and incongruent-first conditions was also tested using the 95% confidence interval (**Figure 4C**). Equivalence here was defined as being within plus or minus 10% the difference [note that the difference RTs is far smaller than actual RTs, so 10% of a difference is miniscule (roughly 12 ms or so)].

RESULTS FOR ACCURACY OF RESPONDING

Because accuracy data are binary at the individual trial level, a generalized estimating equation using a binary logistic equation was used to compare the difference in accuracy between the first two trial blocks in the congruent-first condition and the incongruent-first condition. Accuracy did not differ for any block by ethnicity. Children >7 years made no errors on the button practice that preceded testing on hearts and flowers.

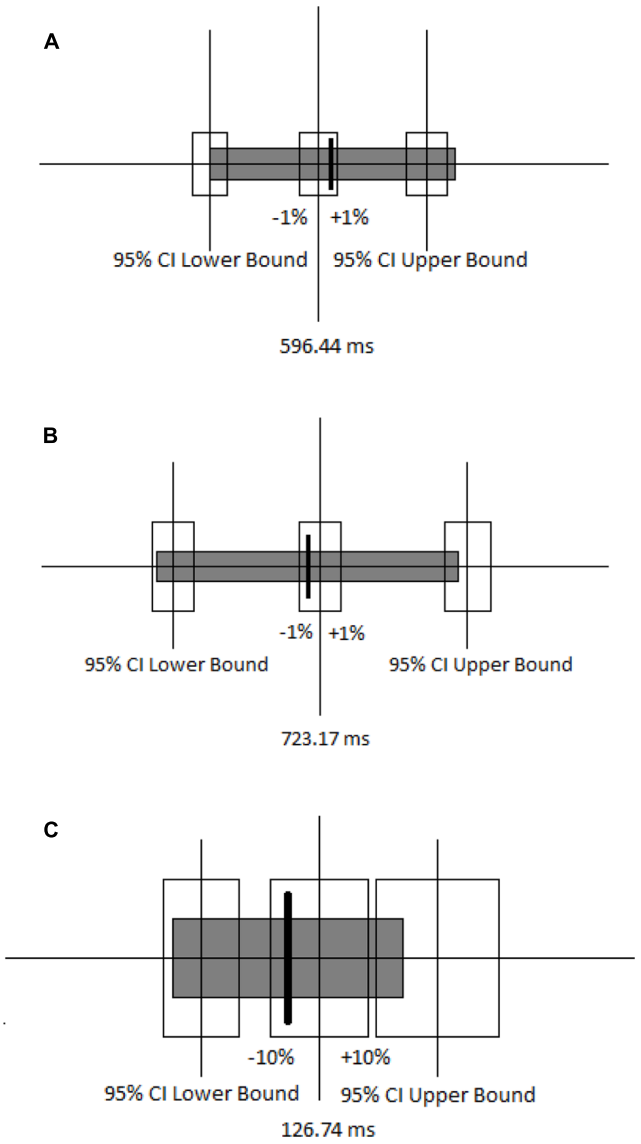


FIGURE 4 | (A) Ninety-five percent confidence interval around the mean RT for the congruent block when it came first is shown by the unfilled boxes. The thick line and gray box indicate the mean and 95% confidence interval for congruent trials when they came second. **(B)** Ninety-five percent confidence interval around the mean RT for the incongruent block when it came first is shown by the unfilled boxes. The thick line and gray box indicate the mean and 95% confidence interval for incongruent trials when they came second. **(C)** Ninety-five percent confidence intervals around the mean RT for the differences between the congruent and incongruent blocks when the congruent block came first is shown by the unfilled boxes. The thick line and gray box indicate the mean and 95% confidence interval for the difference between the congruent and incongruent blocks when the incongruent block came first.

Accuracy improved over age from 6 to 10 years on both the congruent and incongruent blocks (chi square: congruent block: $\chi^2(1, N = 96) = 4.13$, $p < 0.04$, odds ratio = 1.64; incongruent block: $\chi^2(1, N = 96) = 7.11$, $p < 0.01$, odds ratio = 1.39). Excluding 6-year-olds (who were given more time to respond than all other children) accuracy improved over age only on the mixed block [$\chi^2(1, N = 77) = 7.18$, $p < 0.01$]. All children >7 years were correct on all trials in the congruent block. Accuracy did not differ by ethnicity on any block or by gender on the incongruent or mixed blocks. However, girls were correct on more trials than boys in the congruent block (all ages: $\chi^2(1, N = 96) = 5.70$, $p < 0.02$, odds ratio = 2.26; only children 7–10 years old: $\chi^2(1, N = 77) = 9.40$, $p < 0.02$, odds ratio = 3.54).

In the congruent-first condition, participants responded more accurately on the congruent trial block (mean = 97.26%, SD = 5.07%) than on the incongruent one (mean = 92.17%, SD = 8.50%): $\chi^2(1, N = 46) = 7.50$, $p < 0.006$, odds ratio = 4.58. In the incongruent-first condition as well, the percentage of correct responses was higher in the congruent block (mean = 95.50%, SD = 6.77%) than in the incongruent one (mean = 91.00%, SD = 7.96%): $\chi^2(1, N = 50) = 8.23$, $p < 0.004$, odds ratio = 3.02. *At every age*, regardless of the order in which the conditions were tested, children made fewer errors in the congruent block than in the incongruent one (see **Figure 5A**).

An ANOVA within a General Linear Model with age as a continuous between-subject variable, with order of trial blocks and gender as categorical between-subject variables, and with block type (congruent or incongruent) as a categorical within-subject factor was conducted to determine whether the difference in accuracy between the two blocks of trials was similar or different in the two orders of testing. The difference in accuracy between congruent and incongruent blocks did not vary by order of presentation [$F(1,89) = 2.04$, ns]. *At no age* did the within-child difference in accuracy on the two blocks differ significantly by order of presentation (see **Figure 5B**). All of the above also held regardless of ethnicity or gender and there were no significant effects of, or interactions with, gender or ethnicity. Children of 6 years were as accurate on the incongruent block as the congruent one, so including them in the analyses showed a significant increase in the difference in percentage of correct trials on congruent versus incongruent trials by age [$F(1,89) = 5.29, p < 0.02$]. Including only the children who received the same timing parameters (children 7–10 years), there was no change in this difference over age.

Again, a difference that fails to reach significance is insufficient to demonstrate equivalence, so a specified interval of equivalence was again used. The interval was set at 2% because one incorrect answer causes a large change in accuracy values. Both mean accuracy on the congruent block (**Figure 6A**) and on the incongruent block (**Figure 6B**) for the incongruent-first order of testing fell within the interval of equivalence for the congruent-first order of testing. The distribution of percentage of correct responses on incongruent blocks was also equivalent across the two orders of testing (see **Figure 6A**). The distributions of percentage of correct responses on congruent blocks, however, did differ: Children made more errors on congruent trials when they followed incongruent ones than when congruent trials came first, providing a hint of a subtle difference in performance on congruent trials by order of testing. The equivalence of the difference in accuracy between the congruent and incongruent blocks in both congruent-first and incongruent-first conditions was also tested using the 95% confidence interval (**Figure 6C**). This interval of equivalence was set at 1% because a difference score always has

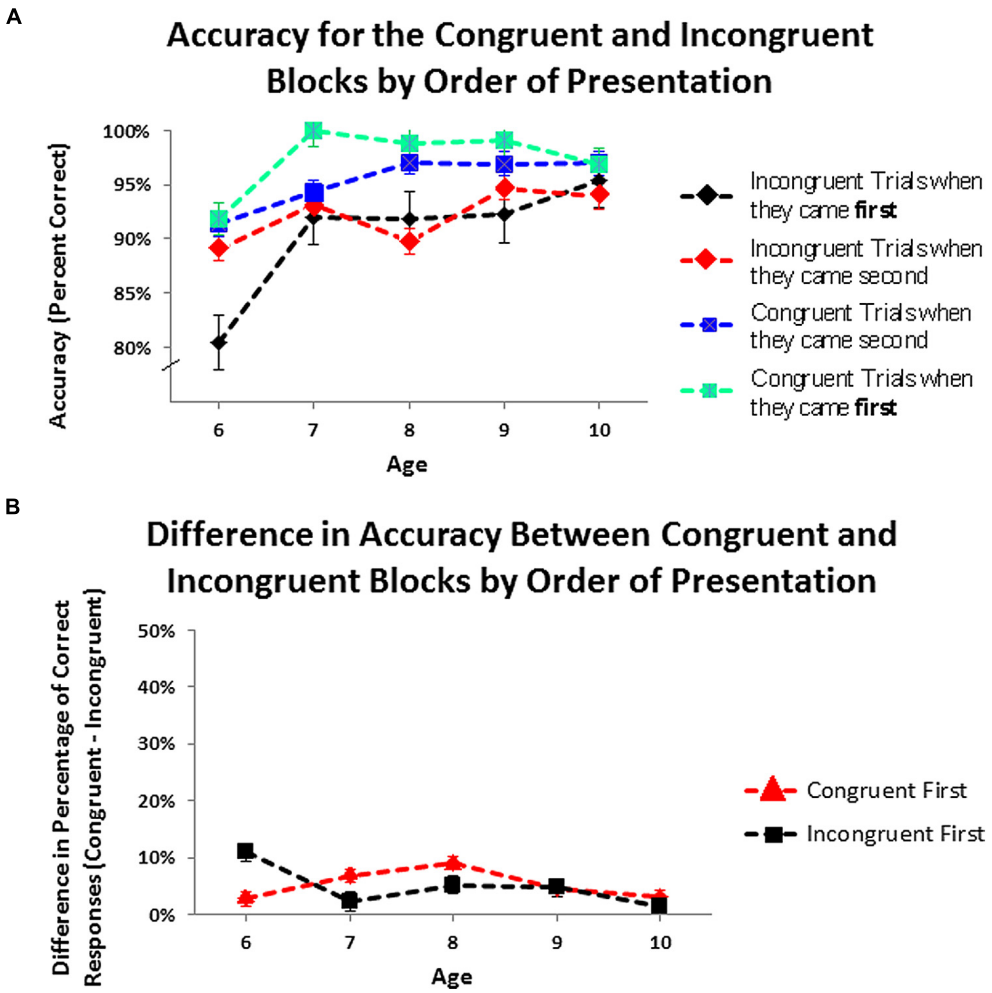


FIGURE 5 | (A) Accuracy on the congruent and incongruent blocks by order of presentation. **(B)** Difference in accuracy between congruent and incongruent blocks by order of presentation.Values greater than zero indicate that children were correct on more trialsin congruent than in incongruent blocks.

a small range of variability. The difference in accuracy between the two blocks was equivalent regardless of which block came first.

DISCUSSION

This study explored what the critical difference is between incongruent and congruent blocks that accounts for why children perform so much worse on incongruent blocks. For the first time we know of, the order in which congruent and incongruent blocks were presented to children was varied. Worse performance on the incongruent block of the hearts and flowers task when it comes second could be accounted for by greater working memory demands (subjects might still be holding the first rule in mind when performing Block 2), greater inhibitory demands, task-switching demands, or some combination of those. However, worse performance on the incongruent block when it comes first (as found here) can be accounted for only by greater demands on inhibition.

Regardless of the order in which the congruent and incongruent blocks were presented, children at every age were slower and made more errors on the incongruent block than the congruent one. That is, they performed worse on the incongruent block even when it was presented first, and this difference in performance was no greater when the incongruent block came second. These results strongly support that the source of the difficulty for children is *not* switching from the rule in Block 1 to the rule in Block 2, nor does the source of their problem seem to be holding in mind the rule for Block 1 when they perform Block 2 (not having efficiently deleted it from working memory); the source of their difficulty seems to be the need to inhibit a pre-potent response on incongruent trials. These results also show that varying the demand on inhibition (the incongruent block requires inhibition of a pre-potent behavioral tendency whereas the congruent block does not) holding working memory demands constant (when the congruent block is presented first it requires holding one rule in mind and when the incongruent block is presented first it, too, requires holding only one rule in mind) is sufficient to produce a decrement in children's performance evident both in poorer speed and accuracy.

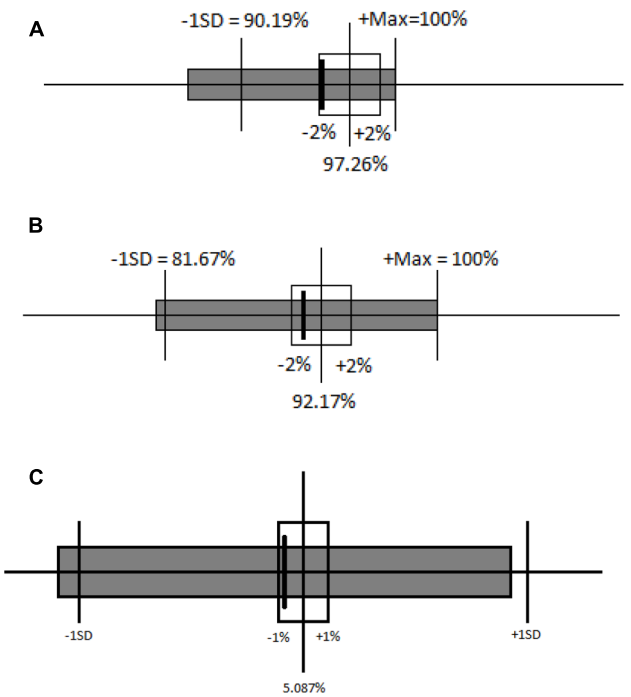


FIGURE 6 | (A) Ninety-five percent confidence interval around the mean accuracy for the congruent block in the congruent-first condition (unfilled boxes and solid lines) and comparison to the incongruent-first condition (solid box), plus or minus 2%. **(B)** Ninety-five percent confidence interval around the mean accuracy for the incongruent block in the congruent-first condition (unfilled boxes and solid lines) and comparison to the incongruent-first condition (solid box), plus or minus 2%. **(C)** Ninety-five percent confidence interval around the mean difference in accuracy between congruent and incongruent blocks in congruent-first and incongruent-first conditions. Solid lines represent the accuracy difference in the congruent-first condition and the gray box represents the accuracy difference in the incongruent-first condition.

A memory theorist might protest that worse performance on incongruent trials even when they come first could result from the difficulty of maintaining the rule sufficiently active in working memory for it to "win" in the battle for controlling behavior in the face of interference from the natural inclination to press on the same side as a stimulus. This seems to allow for no disproof of a working memory hypothesis, however, because it asserts that whenever the demand on inhibition is increased (whenever a strong disposition to act a certain way must be suppressed or overridden) ipso facto the demand on working memory is increased. What we know about the incongruent condition is that a strong competing response is present (a strong tendency to give a response that would be incorrect must not win; it must be inhibited). We also know that the incongruent and congruent conditions require holding and manipulating in working memory only one rule each. Those are objective behavioral observations. It is an unproven hypothesis that inhibition of a competing response is accomplished by working memory "working harder." It is also an unproven hypothesis that inhibition of a competing response is accomplished by executive attention working harder to keep one's attention focused on the relevant rule. This paper reports what is behaviorally available for observation.

These results provide evidence of the consequences of a greater inhibitory demand (on incongruent trials), independent of any difference in the quantity or complexity of what must be held in working memory. In the face of no change in the working memory demand, increasing the demand on inhibitory control alone is sufficient to induce more errors and slower responding in children. Adults may not appreciate how inordinately difficult inhibition is for young children because it is so much less difficult for adults (adults show no difference in performance on congruent and incongruent blocks of the hearts and flowers task, or usually of Simon or spatial Stroop tasks, showing errors and slower responding only on mixed blocks [Lu and Proctor, 1995]). Often conditions differ in both working memory and inhibitory control demands making it impossible to attribute differences in performance specifically to working memory or inhibitory control. Here, where demands on working memory and inhibitory have been dissociated, it is possible to see that increasing inhibitory control demands alone is sufficient to induce worse performance in children 6–10 years of age.

ACKNOWLEDGMENTS

The research reported here was funded by a grant from the National Institute on Drug Abuse (NIDA R01 #DA019685). The senior author was supported by a grant from the National Institute of Mental Health (NIMH R01 #MH 071893) during the writing of this paper. We thank all the children who participated and their parents and teachers for allowing them to participate.

REFERENCES

Allport, A., and Wylie, G. (2000). "Task switching, stimulus-response bindings, and negative priming," in *Control of cognitive processes: Attention and Performance XVII*, eds S. Monsell and J. Driver (Cambridge, MA: MIT Press), 35–70.

Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359

Cepeda, N. J., Kramer, A. F., and Gonzalez, D. S. (2001). Changes in executive control across the life span: examination of task-switching performanc. *Dev. Psychol.* 37, 715–730. doi: 10.1037/0012-1649.37.5.715

Chatham, C. H., Claus, E. D., Kim, A., Curran, T., Banich, M. T., and Munakata, Y. (2012). Cognitive control reflects context monitoring, not motoric stopping in response inhibition. *PLoS ONE* 7:2. doi: 10.1371/journal.pone.0031546

Cohen, J. D., Braver, T. S., and Brown, J. W. (2002). Computational perspectives on dopamine function in prefrontal cortex. *Curr. Opin. Neurobiol.* 12, 223–229. doi: 10.1016/S0959-4388(02)00314-8

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–607. doi: 10.1038/386604a0

Crone, E. A., Bunge, S. A., and Van Der Molen, M. W. (2006). Switching between tasks and responses: a developmental study. *Dev. Sci.* 9, 278–287. doi: 10.1111/j.1467-7687.2006.00490.x

Davidson, M. C., Amso, A., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 year: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006

DeSoto, M. C., Fabiani, M., Geary, D. C., and Gratton, G. (2001). When in doubt, do it both ways: brain evidence of the simultaneous activation of conflicting motor responses in a spatial Stroop task. *J. Cogn. Neurosci.* 13, 523–536. doi: 10.1162/08989290152001934

D'Esposito, M., Aguirre, G. K., Zarahn, E., Ballard, D., Shin, R. K., and Lease, J. (1998). Functional MRI studies of spatial and nonspatial working memory. *Cogn. Brain Res.* 7, 1–13. doi: 10.1016/S0926-6410(98)00004-4

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., and Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature* 378, 279–281. doi: 10.1038/378279a0

Diamond, A. (2009). When in competition against engrained habits, is conscious representation sufficient or is inhibition of the habit also needed? *Dev. Sci.* 12, 20–22. doi: 10.1111/j.1467-7687.2008.00773.x

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Diamond, A., Barnett, S., Thomas, J., and Munro, S. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Diamond, A., Briand, L., Fossella, J., and Gehlbach, L. (2004). Genetic and neurochemical modulation of prefrontal cognitive functions in children. *Am. J. Psychiatry* 161, 125–132. doi: 10.1176/appi.ajp.161.1.125

Edgin, J. O., Mason, G. M., Allman, M. J., Capone, G. T., DeLeon, I., Maslen, C., et al. (2010). Development and validation of the Arizona Cognitive Test Battery for Down syndrome. *J. Neurodev. Disord.* 2, 149–164. doi: 10.1007/s11689-010-9054-3

Egner, T., and Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nat. Neurosci.* 8, 1784–1790. doi: 10.1038/nn1594

Evans, J. W., Fossella, J., Hampson, E., Kirschbaum, C., and Diamond, A. (2009). Gender differences in the cognitive functions sensitive to the level of dopamine in prefrontal cortex. *Poster presentation, Inaugural Conference on "Executive Function and Dysfunction," (Jan. 15) University of Boulder, Boulder, CO and at the Association for Psychological Science Annual Meeting (May 25)*, San Francisco, CA.

Fitts, P. M., and Seeger, C. M. (1953). S-R compatibility: spatial characteristics of stimulus and response codes. *J. Exp. Psychol.* 81, 174–176. doi: 10.1037/h0062827

Gazzaley, A., Cooney, J. W., Rissman, J., and D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nat. Neurosci.* 8, 1298–1300. doi: 10.1038/nn1543

Georgopoulos, A. (1994). Population activity in the control of movement. *Int. Rev. Neurobiol.* 37, 103–119.

Georgopoulos, A. P., Lurito, J. T., Petrides, M., Schwartz, A. B., and Massey, J. T. (1989). Mental rotation of the neuronal population vector. *Science* 243, 234–236. doi: 10.1126/science.2911737

Gerardi-Coulton, G. (2000). Sensitivity to spatial conflict and the development of self-regulation in children 24-36 months of age. *Dev. Sci.* 3, 397–404. doi: 10.1111/1467-7687.00134

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychol. Sci.* 4, 294–298. doi: 10.1111/j.1467-9280.1993.tb00567.x

Hanania, R., and Smith, L. B. (2010). Selective attention and attention switching: towards a unified developmental approach. *Dev. Sci.* 13, 622–635. doi: 10.1111/j.1467-7687.2009.00921.x

Hartman, M., and Hasher, L. (1991). Aging and suppression: memory for previously relevant information. *Psychol. Aging* 6, 587–594. doi: 10.1037/0882-7974.6.4.587

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychol.* 136, 189–202. doi: 10.1016/j.actpsy.2010.04.011

Hommel, B., Proctor, R. W., and Vu, K.-P. L. (2004). A feature integration account of sequential effects in the Simon task. *Psychol. Res.* 68, 1–17. doi: 10.1007/s00426-003-0132-y

Kornblum, S., Stevens, G. T., Whipple, A., and Requin, J. (1999). The effects of irrelevant stimuli: 1. The time course of S-S and S-R consistency effects with Stroop-like stimuli, Simon-like tasks, and their factorial combinations. *J. Exp. Psychol.* 25(A), 688–714. doi: 10.1037/0096-1523.25.3.688

Kunde, W., and Stocker, C. (2002). A Simon effect for stimulus-response duration. *Q. J. Exp. Psychol.* 55A, 581–592. doi: 10.1080/02724980143000433

Leroux, G., Joliot, M., Dubal, S., Mazoyer, B., Tzourio-Mazoyer, N., and Houde, O. (2006). Cognitive inhibition of number/length interference in a Piaget-like task in young adults: evidence from ERPs and fMRI. *Hum. Brain Mapp.* 27, 498–509. doi: 10.1016/j.clinph.2009.06.003

Lesaffre, E. (2008). Superiority, equivalence, and non-inferiority trials. *Bull. NYU Hosp. Jt. Dis.* 2, 150–154.

Levy, B. J., and Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends Cogn. Sci. (Regul. Ed.)* 6, 299–305. doi: 10.1016/S1364-6613(02)01923-X

Lu, C. H., and Proctor, R. W. (1995). The influence of irrelevant location information on performance: a review of the Simon and spatial Stroop effects. *Psychon. Bull. Rev.* 2, 174–207. doi: 10.3758/BF03210959

Monsell, S., and Driver, J. (eds). (2000). *Control of Cognitive Processes: Attention and Performance XVIII.* Cambridge, MA: MIT Press.

Mullane, J. C., Corkum, P. V., Klein, R. M., and McLaughlin, E. (2009). Interference control in children with and without ADHD: a systematic review of Flanker and Simon Task Performance. *Child Neuropsychol.* 15, 321–342. doi: 10.1080/09297040802348028

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., and O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends Cogn. Sci. (Regul. Ed.)* 15, 453–459. doi: 10.1016/j.tics.2011.07.011

Nieuwenhuis, S., and Yeung, N. (2005). Neural mechanisms of attention and control: losing our inhibitions? *Nat. Neurosci.* 8, 1631–1633. doi: 10.1038/nn1205-1631

Owen, A. M., Evans, A. C., and Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within lateral frontal cortex: a positron emission tomography study. *Cereb. Cortex* 6, 31–38. doi: 10.1093/cercor/6.1.31

Petrides, M. (1994). "Frontal lobes and working memory: evidence from investigations of the effects of cortical excisions in nonhuman primates," in *Handbook of Neuropsychology*, Vol. 9, eds F. Boller and J. Grafman (Amsterdam: Elsevier Science Publishers) 59–82.

Petrides, M. (1995). Functional organization of the human frontal cortex for mnemonic processing: evidence from neuroimaging studies. *Ann. N. Y. Acad. Sci.* 769, 85–96. doi: 10.1111/j.1749-6632.1995.tb38133.x

Simon, R. J., and Rudell, A. P. (1967). Auditory S-R compatibility: the effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304. doi: 10.1037/h0020586

Smith, E. E., and Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science* 283, 1657–1661. doi: 10.1126/science.283.5408.1657

Smith, E. E., Jonides, J., Marshuetz, C., and Koeppe, R. A. (1998). Components of verbal working memory: evidence from neuroimaging. *Proc. Natl. Acad. Sci. U.S.A.* 95, 876–882. doi: 10.1073/pnas.95.3.876

Valle-Inclan, F. (1996). The Simon effect and its reversal studied with event-related potentials. *Int. J. Psychophysiol.* 23, 41–53. doi: 10.1016/0167-8760(96)00027-X

Yeung, N., Nystrom, L. E., Aronson, J. A., and Cohen, J. D. (2006). Between-task competition and cognitive control in task switching. *J. Neurosci.* 26, 1429–1438. doi: 10.1523/JNEUROSCI.3109-05.2006

Zanto, T. P., and Gazzaley, A. (2009). Neural suppression of irrelevant information underlies optimal working memory performance. *J. Neurosci.* 29, 3059–3066. doi: 10.1523/JNEUROSCI.4621-08.2009

Zaitchik, D., Iqbal, Y., and Carey, S. (2013). The effect of executive function on biological reasoning in young children: an individual differences study. *Child Dev.* 85, 160–175. doi: 10.1111/cdev.1214

Zelazo, P. D., Muller, U., Frye, D., and Marcovitch, S. (2003). The development of executive function in early childhood. *Monogr. Soc. Res. Child Dev.* 68, 137. doi: 10.1111/j.0037-976X.2003.00269.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 December 2013; accepted: 24 February 2014; published online: 14 March 2014.

Citation: Wright A and Diamond A (2014) An effect of inhibitory load in children while keeping working memory load constant. Front. Psychol. ***5****:213. doi: 10.3389/fpsyg.2014.00213*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Wright and Diamond. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 29 July 2014
doi: 10.3389/fpsyg.2014.00831

Age-related changes in the temporal dynamics of executive control: a study in 5- and 6-year-old children

Joanna Lucenet* and Agnès Blaye

CNRS, Laboratoire de Psychologie Cognitive, UMR 7290, Aix-Marseille Université, Marseille, France

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Yuko Munakata, University of Colorado, USA
Jason F. Reimer, California State University, USA
Katharine A. Blackwell, Salem College, USA

****Correspondence:***
Joanna Lucenet, CNRS, Laboratoire de Psychologie Cognitive, UMR 7290, Aix-Marseille Université, Pôle 3C, 3 Place Victor Hugo, 13331 Marseille, France
e-mail: joanna.lucenet@gmail.com

Based on the Dual Mechanisms of Control theory (Braver et al., 2007), this study conducted in 5- and 6-year-olds, tested for a possible shift between two modes of control, proactive vs. reactive, which differ in the way goal information is retrieved and maintained in working memory. To this end, we developed a children-adapted version of the AX-Continuous-Performance Task (AX-CPT). Twenty-nine 5-year-olds and 28-6-year-olds performed the task in both low and high working-memory load conditions (corresponding, respectively, to a short and a long cue-probe delay). Analyses suggested that a qualitative change in the mode of control occurs within the 5-year-old group. However, quantitative, more graded changes were also observed both within the 5-year-olds, and between 5 and 6 years of age. These graded changes demonstrated an increasing efficiency in proactive control with age. The increase in working memory load did not impact the type of dynamics of control, but had a detrimental effect on sensitivity to cue information. These findings highlight that the development of the temporal dynamics of control can be characterized by a shift from reactive to proactive control together with a more protracted and gradual improvement in the efficiency of proactive control. Moreover, the question of whether the observed shift in the mode of control is task dependant is debated.

Keywords: goal setting, reactive control, proactive control, context processing, cognitive development, executive functions

INTRODUCTION

Executive control, defined as the ability to regulate, coordinate, and guide one's thoughts and behaviors toward goals, is probably one of the most critical aspects of human cognition. Indeed, executive control is involved in the development of many cognitive and social skills during childhood such as language (Deak, 2003), theory of mind (Carlson and Moses, 2001; for a review see Miller and Marcovitch, 2012), reading, reasoning and arithmetic (Blair and Razza, 2007; Clark et al., 2013), and emotion regulation (Carlson and Wang, 2007; Eisenberg and Sulik, 2012). It is now well accepted that executive control dramatically develops between the ages of 3 and 6 years (Wright et al., 2003; Carlson et al., 2004; Chevalier et al., 2012). Although these developmental changes have been viewed as resulting merely from an increase in the efficiency of control, recent work suggests that age-related qualitative differences in the control strategies used may also contribute to this development (Chatham et al., 2009; Dauvier et al., 2012; Chevalier et al., 2013). The aim of the present study was to assess whether qualitative changes in the mode of control might occur between the ages of 5 and 6. Specifically, we investigated potential age-related differences in the use of two modes of control proactive vs. reactive which differ in terms of the activation and maintenance of goal representations. To this end, 5- and 6-year-old children were presented with an adapted version of the AX-Continuous Performance Task (Braver et al., 2001).

Executive control is traditionally viewed as composed of three functions: inhibition, flexibility, and working memory updating (Miyake et al., 2000; Lehto et al., 2003; Carlson, 2005). Despite their partial independence, there is converging evidence that these functions share a common base (Miyake et al., 2000; Friedman et al., 2008; Miyake and Friedman, 2012). These authors have proposed that active maintenance of a goal representation and its use to bias task processing under conditions of interference could account for this common core component. Recent empirical work supports this hypothesis and suggests that the activation and maintenance of task–goal information may play a critical role in efficient control, both in adults (Baddeley et al., 2001; Rubinstein et al., 2001; Emerson and Miyake, 2003; Gruber and Goschke, 2004) and in children (Morton and Munakata, 2002; Zelazo et al., 2003; Towse et al., 2007; Chevalier and Blaye, 2009; Chevalier et al., 2010; Blaye and Chevalier, 2011). Developmental studies reveal that the representation and active maintenance of task–goal information improve from childhood to adulthood (Karbach and Kray, 2007; Chevalier and Blaye, 2009; Chevalier et al., 2010, 2012).

Preschool-aged children's poor flexibility has recently been shown to depend, at least in part, on failures of goal maintenance (Marcovitch et al., 2007, 2010) and in goal representation (Chevalier and Blaye, 2009). Marcovitch et al. (2007, 2010) used a variant of the Dimensional Change Card Sorting task (DCCS; Zelazo et al., 1996). The DCCS task consists in matching cards depicting bidimensional objects (e.g., red rabbits and blue boats) to one of two target cards. In a first block of trials, children are required to match cards according to one dimension (e.g., shape); In the second block (post-switch), they are required to sort stimuli

according to the other dimension (here, color). Studies on this task in young children have shown that they succeed in sorting cards according to the first dimension, but fail to switch to the second rule after sorting by the first, and perseverate to match stimuli following the first rule. Marcovitch et al. (2007, 2010) tested the hypothesis that failure in the post-switch block was due to a flaw in maintenance of the goal, here of the matching rules, by manipulating the frequency of "conflict" cards in the post-switch block. Conflict cards require opposite matching depending on the rule to be used (because they match one target card on one dimension and the other on the other dimension). A high proportion of these cards thus lead to a greater need for goal maintenance, whereas a low proportion, involving many no-conflict cards that can be sorted independently of the rule to be applied, makes goal maintenance more demanding. As expected, preschoolers' performance was worse when the frequency of conflict cards was low. Hence, despite understanding task instructions, young participants may fail to execute them effectively, a phenomenon that is referred to as "goal neglect" (Duncan et al., 1996). Chevalier and Blaye (2009) investigated the critical role of the activation of a task goal representation by manipulating task cues in a task-switching paradigm requiring participants to switch between shape- and color-matching rules. The authors graded the transparency of task-cues (i.e., the degree of association between cues and goals) and found that arbitrary cues made it more difficult for 5- and 6-year-old children to activate a representation of what to do next. Interestingly, the effect of cue transparency decreased in older children and adults, thereby suggesting that preschoolers' struggle to translate arbitrary cues into task goals might reflect lower flexibility in comparison to older children. The nature of the changes contributing to the development of both the activation and maintenance of goal representations remains to be explored.

Recent research has evidenced age-related qualitative changes in control strategies that might promote the development of cognitive flexibility (Chevalier et al., 2011, 2013) and working memory (Camos and Barrouillet, 2011) from preschool to school ages. For instance, Camos and Barrouillet (2011) observed changes from a strategy of passive maintenance of memoranda in preschoolers, to a strategy of refreshing in school-age children. Using a flexibility task, Chevalier et al. (2013) produced the first findings suggesting a difference between 5 year-olds and 10 year-olds in goal representation and maintenance strategies. In addition to the task cues that indicated which task to perform next, as in the traditional task-switching paradigm, they provided transition cues specifying the nature of the transition between two consecutive trials: task repetition vs. task alternation. These transition cues were helpful for the younger participants, but proved to be detrimental to 10-year-olds' flexibility scores, thereby suggesting that the two age groups employed different strategies in task–cue processing, and hence in goal representation. In the present paper, we further explore the nature of the changes that underlie developmental improvements in children's ability to activate and maintain goal representations. Although developed to account for adult control, the Dual Mechanisms of Control theory (DMC theory, Braver et al., 2001, 2007; Braver and Barch, 2002; Braver, 2012) offers a theoretical framework for examining this question. Specifically, this approach offers an account of the way individuals retrieve and maintain goal-related information, and use it to guide processing (Braver, 2012). The DMC theory makes a qualitative distinction between two modes of control engaged under conditions of interference. It is noteworthy that interference can be induced by either irrelevant stimulus information or irrelevant dominant responses. These two control modes, respectively, called "proactive" and "reactive" have different temporal dynamics and neural substrates. The use of a proactive mode of control involves not only the retrieval of a representation of the goal in advance of the stimuli requiring a response, but also the active maintenance of this representation in working memory in order to bias processing towards task-relevant information. In contrast, with a reactive form of control, the goal is retrieved "just in time," after the occurrence of the stimulus and its representation is transiently maintained in working memory.

Empirically, the two forms of control are assessed using the AX-Continuous Performance Task (AX-CPT, Braver et al., 2001). In this paradigm, cue–probe pairs are presented sequentially. Participants have to give a target vs. non-target response to each probe stimulus based on the cue stimulus presented immediately before it. In adults and older children, letters are used as cues and probes. At probe onset, participants are required to press one of two response keys, associated to either target or non-target responses. A target response is required when an A cue is followed by an X probe (AX target trials, thus the name AX-CPT), whereas non-target responses are to be given for all other cue–probe pairs (AY, BX, and BY trials, where Y and B represent any letters other than A and X). AX trials make up 70% of trials, while the frequency of each of the other three types of non-target trials is 10. Formally, since both AY and BX trials involve one letter that is strongly associated with the target response whereas a non-target response is expected, they could be considered as both requiring inhibition and then, could lead to similar performance (e.g., Paxton et al., 2008). The point of analyzing AX-CPT performance, however, is to reveal the pattern of differences between these two trial types. This pattern is considered as an index of the degree to which participants' attention is drawn to the cue. Participants who use a proactive form of control engage in active preparation of their response to the probe when they see the cue. Hence, as the high proportion of AX trials creates a strong expectancy to give a target response it is detrimental to performance when the A cue appears and it is not followed by an X probe (i.e., AY trials). Indeed, this situation is specifically costly in terms of inhibition because participants have to reject the tendency to give a target response to the Y probe. The high AX trials' frequency also induces a bias to produce a target response when an X probe is not preceded by an A cue (i.e., BX trials). Therefore, responding correctly to BX trials requires participants to actively maintain the B cue: because orienting attention towards B cue through active maintenance has the effect of inhibiting goal-irrelevant information, it aids participants to reject the strong tendency to give a target response to the X probe. The reverse pattern is expected in participants who have difficulty using goal-related information (i.e., who exercise reactive control): they do not anticipate their response to the

probe according to the cue and make their decision only after the probe display. Because participants using reactive control do not actively maintain the cue during the cue–probe delay, they do not need to overcome the strong bias that an A cue is followed by an X probe. Hence, the use of reactive control should lead to higher performance on AY trials. By contrast, in order to produce a correct non-target response to X probes which follow an invalid cue (i.e., BX trials), participants have to retrieve the cue that they did not actively maintain in order to inhibit their tendency to give a target response when seeing the X probe. In sum, proactive control is typically evidenced by better performance on BX trials than on AY trials, while reactive control is reflected by better performance on AY than BX trials. It should be noted that performance on BY trials is not expected to differ between proactive and reactive participants, as neither the cue nor the probe is associated to a target response on this kind of trial.

Data from studies in adults investigating the relations between mode of control and working-memory on the one hand, and form of control and neural substrates on the other hand, suggest converging predictions on what could be the development of the dynamics of control. There is empirical evidence that working memory capacity plays a role in the temporal dynamics of control in adults. For instance, Redick (2014) showed in young adults that individuals with high working memory capacity tend to use a proactive form of control more often than individuals with low working memory capacity. The increase in working memory capacity over childhood (Gathercole et al., 2004) suggests one reason why younger children should encounter more difficulties using proactive control than older ones. Moreover, according to the DMC theory, proactive control is subserved by a phasic signal from the dopaminergic (DA) system prior to stimulus onset and by sustained activation of the lateral prefrontal cortex (PFC), a region that is known to be involved in the active maintenance of goal-related information. By contrast, reactive control does not involve a burst of DA activity, but instead involves transient activation of the lateral PFC when triggered by critical stimuli. In this case, the reactivation of goal-related information requires either the detection of interference through additional conflict monitoring regions such as the anterior cingular cortex (ACC), or the retrieval of associations through temporal or cortical brain areas. Given that the frontal lobes are known to be the last brain regions to develop, reaching maturity only in adolescence (Casey et al., 2000), researchers have hypothesized that younger children's less efficient executive control might be related to the lesser powers of proactive control resulting from the immaturity of their frontal cortex (Braver, 2012; Munakata et al., 2012).

It is noteworthy that the developmental course of working memory and neural deterioration in aging suggests symmetrical developmental predictions. These predictions have received some empirical support: a shift from a proactive to a reactive mode of control with aging has been observed using both behavioral and neurophysiological measures (Braver et al., 2001, 2005; Paxton et al., 2008). The dynamics of control during childhood, by contrast, remain under-investigated, with only two studies addressing this question in children older than 8 (Lorsbach and Reimer, 2008, 2010) and only one contrasting 3.5- and 8-year-olds (Chatham et al., 2009). Lorsbach and Reimer (2010) found that children between the ages of 9 and 11 already engaged a proactive form of control. They observed a developmental increase in efficiency of this form of control only for a long cue–probe interval, suggesting that goal maintenance mechanisms are involved in the development of executive control during childhood. Chatham et al. (2009) drew similar conclusions in younger children. The authors used an adapted version of the AX-CPT paradigm with pictures instead of letters. Pupillometry measures and behavioral observations both revealed that 8-year-olds children engaged in intense mental efforts during the cue–probe interval, thereby suggesting that they struggled to actively maintain the cue in working memory. Younger children (3.5 years old) did not show any maintenance-related effort during this interval, but instead showed a reactive peak during probe display on BX trials. Although these data suggest a shift from reactive to proactive control during childhood, the turning point of these qualitative changes is unclear due to the large age gap (i.e., 3.5- vs. 8-year-olds) between the two groups. Moreover, the task used differed from the standard AX-CPT task in ways that might affect interpretations of the patterns of behavior. Not only did the task involve only two cues and two probes instead of the great diversity of letters referred to as B and Y in the standard AX-CPT, but also, in contrast to the arbitrariness of the cue–probe associations in the standard task, here it was contextualized in a story (e.g., As Spongebob (A) likes watermelon (X), a press on happy face is expected when Spongebob appears followed by the watermelon). In sum, it is unclear whether performances on this task are directly comparable to those obtained with the standard AX-CPT. Hence, further data using a task closer to the standard one is then required to enable a comparison between performance in young children and data previously obtained on older ones. In investigating a narrower age range, we expected to pinpoint the qualitative shift form reactive to proactive control.

In light of the finding of a substantial improvement in the ability to retrieve and maintain goal-related information in working memory between the ages of 5 and 6 years (Chevalier and Blaye, 2009; Chevalier et al., 2010; Camos and Barrouillet, 2011), we selected this age range to explore a potential shift from reactive to proactive control in children. Investigating this age range seems also particularly relevant with respect to Blackwell and Munakata's (2013) interpretation of performance of 6-year-old children in a flexibility task, as revealing a shift from a reactive to a proactive mode of control. Moreover, since it has been suggested that working memory is critical in determining mode of control (Redick, 2014), we assessed these developmental differences in two different working memory load conditions by varying the length of the cue–probe delay. We used a new child-specific version of the AX-CPT, designed to be as similar as possible to the adult version of the task. Because this study is the first to investigate the dynamics of control in the age range of 5–6 years, alternative predictions can be made. First, there could be a shift from reactive to proactive control, which would then be evidenced by the typical pattern of reactive control in the younger group, with better performance on AY trials than on BX trials, and the reverse pattern in the older, proactive group. Second, it is also plausible that both age groups

already use proactive control: in this case, changes between the ages of 5 and 6 would be evidenced by greater efficiency in retrieving and actively maintaining goal-related information. This should be reflected by increased difficulties in inhibiting a target response to the Y probe when presented following the A cue, and/or better ability to anticipate the need for a non-target response when presented along with a B cue: Third, both age groups could perform the task using a reactive mode of control. In this case, children may have difficulty anticipating the need for a non-target response when B cue is presented, but perform better when an A cue is followed by a non-target Y probe. If two profiles (proactive vs. reactive) would be observed, we hypothesized that children using a proactive control should demonstrate higher speed of processing, especially in the more demanding situation. This assumption was based on research on working memory that showed that participants associated with greater memory span (i.e., those better able to maintain information) are the faster ones (e.g. Barrouillet et al., 2009).

Following Lorsbach and Reimer's (2010) observations in older children, we expected the differences between the two age groups to increase under conditions of high working memory load (long cue–probe delay). Finally, in order to track quantitative changes, we used an index of context sensitivity (d') susceptible to provide a more graded picture of the extent to which children make target response to the X probe according to the cue presented ahead. One may hypothesize that sensitivity to cue information increases from the age of 5 to 6. As sensitivity to cue information can rely on proactive maintenance or reactive retrieval of the cue to guide response to X probe, we hypothesize a reduction of this sensitivity when the cue–probe delay increases because the high working-memory load in this case may hinder cue maintenance.

MATERIALS AND METHODS

PARTICIPANTS

Twenty-nine 5-year-olds (M = 5.80, SD = 0.26; 60% female) and twenty-eight 6-year-olds (M = 6.70, SD = 0.24; 56% female) were recruited from two French preschools and two French primary schools. Parental consent was given for all children, and the experiment was administered individually in a quiet room at the school. Most children were Caucasian and came from middle-class backgrounds, although no data were collected on race and socioeconomic status. Two additional preschoolers and one first-grader also began the experiment but were excluded from analyses because they were disturbed by an unexpected event in the room or they decided to stop the task while in progress.

MATERIALS AND PROCEDURE

We created a child-adapted AX-CPT, replacing the letter stimuli from the original task with black-and-white drawings of animals. The animals were chosen on the basis of identification and naming norms established in 5-year-olds (Chalard et al., 2003; Cannard et al., 2005). As children performed the task twice for each delay condition, two different sets of 13 black-and-white drawings of animals were used in order to prevent boredom. The use of the two sets was counterbalanced across the two conditions. As for letters used in the classic version of the AX-CPT, the animals used for target trials in each set (A cues and X probes) were randomly chosen among each animal list and maintained constant for all participants[1].

Before performing the task, we made sure that all the participants could name each of the animals used as stimuli. AY and BX non-target trials consisted in 12 possible combinations of animal pairs, and BY non-target trials consisted in 132 possible combinations of animal pairs. Task instructions were provided to children as follows: "You will see animals on the screen; these animals run in pairs, one after the other ("ces animaux courent deux par deux, l'un après l'autre")." In one set of animals, children were given the following instruction: "when you first see the hen (A cue) and then the cat (X probe), press the green button, otherwise press the red one." For the other set of animals, they were told "when you first see the frog (A cue) and then the donkey (X probe), press the green button, otherwise press the red one." Children were instructed to respond as quickly and accurately as possible. To ensure that they had memorized the instructions, they were twice shown 4 pairs of sheets of paper mimicking four successive displays of cue and probe combinations on the screen (i.e., AX, AY, BX, BY), once before moving on to the computer training, and once at the end of each session. For each pair, children were questioned about the correct response button to press and were asked to justify their answer to test whether they remembered the rule. All children succeeded in recalling the instructions (showing the correct response button and justifying their response by recounting the rules).

Children were tested individually in two cue–probe delay conditions (1500 ms for the short delay vs. 5500 ms for the long delay) in a counterbalanced order across participants, distributed into two sessions lasting approximately 20–30 min each. A 30-min break was given between the two conditions, during which participants returned to their classroom. Pictures were presented sequentially on a HP Compaq 9000 laptop, using the E-Prime software (Psychology Software Tools, Inc., 2007). Each trial began with the presentation of a fixation cross on the screen for 1500 ms. A cue was then presented at the center of the screen for 500 ms (the first animal, A or B, in the cue–probe pair), followed by a blank screen displayed according to the cue–probe delay (short or long). After this delay, a probe appeared at the center of the screen for 500 ms (the second animal, X or Y, in the cue–probe pair; see **Figure 1**). All probes were framed by a fine black line in order to help children differentiate between cues and probes and decide unambiguously when a response was expected. To encourage children to respond quickly, a warning tone was played when responses exceeded a 1500 ms time limit. Seventy percent of trial were AX target trials, and each of the three kinds of non-target trials (AY, BX, and BY) each made up 10% of trials. The pairs of pictures were presented pseudo-randomly; the number of AX trials in a row never exceeded four. Each delay condition involved a training phase followed by an experimental phase. The training phase included three blocks of 20 trials (14 AX trials, two AY, two BX, and two BY) and the testing phase included four blocks of 30

[1]One set of animals included a hen (A cue), a cat (X probe), a giraffe, a mouse, a crocodile, a horse, a cow, a sheep, a snake, a fish, a rabbit, a pig, and a lion. The other set of animals consisted in a donkey (A cue), a frog (X probe), a squirrel, a dolphin, a bee, a duck, a kangaroo, a rooster, a spider, a turtle, a monkey, a dog, and an elephant.

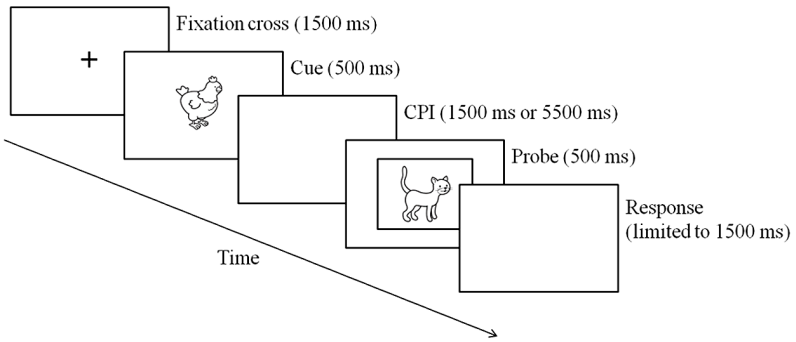


FIGURE 1 | Example of an AX trial sequence.

trials (21 AX trials, three AY, three BX, and three BY), yielding a total of 180 trials.

RESULTS

The main effects of condition order and animal set were not significant, and these variables did not interact significantly with other variables of interest (all $p > 0.10$): they thus were not included in further analyses. Following Lorsbach and Reimer (2008, 2010), we computed different sets of analyses on target trials (AX) and non-target trials (AY, BX, and BY), because they do not involve the same number of trials. The RT on each correct trial was then standardized by subtracting the participant's overall mean to each correct RT and dividing the difference by the same participant's SD. Mean *z*-scores were then calculated for each participant in each condition: negative *z*-scores reveal fast RTs whereas positive *z*-scores correspond to slow RTs. This standardization corrects for individual differences in speed of processing. For clarity, **Table 1** presents a summary of error rates, correct response times and mean *z*-scores. Importantly, because the reliability of error rates is often higher than that of RTs in preschoolers (e.g., Chevalier and Blaye, 2009), analyses on error rates are reported first.

AX TARGET TRIALS

Two similar analyses of variance were run on error rates and mean *z*-scores, with age group (5-year-olds vs. 6-year-olds) as a between-subjects variable and delay (1500 ms short vs. 5500 ms long) as a within-subjects variable.

Age was found to have a significant main effect on error rates, $F(1,55) = 4.66$, $p < 0.05$, $\eta_p^2 = 0.07$, indicating more errors in 5-year-olds ($M = 7.4\%$) than in 6-year-olds ($M = 4.8\%$), but not on *z*-scores, $F(1,55) = 1.74$, $p = 0.18$. The results also revealed a main effect of delay, both on error rates, $F(1,55) = 13.01$, $p < 0.001$, $\eta_p^2 = 0.19$, with higher error rates at the long delay ($M = 7.8\%$) than at the short one ($M = 4.4\%$), and on *z*-scores, $F(1,55) = 227.72$, $p < 0.001$, $\eta_p^2 = 0.80$, indicating faster response times for the short delay ($M = -0.17$) than for the long delay ($M = 0.88$ ms). However, the Age × Delay interaction was not significant, either for error rates or for *z*-scores, $F(1,55) = 0.12$, $p = 0.72$, and $F(1,55) = 1.48$, $p = 0.22$, respectively.

To summarize, error rates on AX trials significantly decreased with age, while latencies on correct trials remained stable between the two age groups. Furthermore, longer delays had a detrimental effect on accuracy and latencies on AX trials.

AY, BX, AND BY NON-TARGET TRIALS

Two analyses of variance were run, following the same design for both error rates and *z*-scores. They involved age group (5-year-olds vs. 6-year-olds) as a between-subjects variable and delay (1500 ms, short vs. 5500 ms, long) and trial type (three types: AY, BX, BY) as within-subjects variables. Because two 5-year-olds and four 6-year-olds produced wrong responses to all trials of one type (i.e., all AY or all BX trials) in the long delay condition, their *z*-score for this type of trial was replaced by the mean *z*-score for their age group to increase statistical power.

A main effect of age was observed on error rates only, $F(1,55) = 5.07$, $p < 0.05$, $\eta_p^2 = 0.08$, revealing that 5-year-olds committed more errors ($M = 23.2\%$) than 6-year-olds ($M = 15.7\%$).Trial type had a significant effect on both performance measures, $F(2,110) = 33.22$, $p < 0.001$, $\eta_p^2 = 0.37$, for error rates and $F(2,110) = 121.46$, $p < 0.001$, $\eta_p^2 = 0.68$, for *z*-scores. Children committed more AY errors ($M = 31.9\%$) than BX errors ($M = 17.8\%$) thereby revealing their use of a proactive mode of control. In addition, BY trials ($M = 8.7\%$) led to fewer errors than AY, $F(1,55) = 55.41$, $p < 0.001$, $\eta_p^2 = 0.50$, and BX trials, $F(1,55) = 33.58$, $p < 0.001$, $\eta_p^2 = 0.37$. Turning to *z*-scores, planned comparisons indicated that latencies were longer on AY trials ($M = 0.77$) than on BX ($M = -0.15$), $F(1,55) = 120.08$, $p < 0.001$, $\eta_p^2 = 0.68$, and BY trials ($M = -0.14$), $F(1,55) = 211.34$, $p < 0.001$, $\eta_p^2 = 0.79$, whereas the latter two did not differ, $F(1,55) = 0.04$, $p = 0.84$. Analyses of response time patterns thus confirmed the above conclusion on error rates. The results also revealed a main effect of delay on error rates, $F(1,55) = 8.59$, $p < 0.01$, $\eta_p^2 = 0.13$, revealing higher error rates at the long delay ($M = 22.3\%$) compared to the short delay ($M = 16.6\%$). A main effect of delay on *z*-scores was also observed, $F(1,55) = 4.14$, $p < 0.05$, $\eta_p^2 = 0.07$, with shorter latencies at the short delay ($M = 0.07$) than the long delay ($M = 0.23$).

Turning to interactions for both measures, only two interactions revealed significant. The interaction between age and trial type was significant on error rates, $F(2,110) = 4.86$, $p < 0.01$, $\eta_p^2 = 0.08$, and on *z*-scores, $F(2,110) = 5.15$, $p < 0.01$, $\eta_p^2 = 0.08$. A Delay × Trial Type interaction was obtained both on error

Table 1 | Mean error rates, correct RTs and *z*-scores by age group and trial type.

	Trial condition							
	Short cue-probe delay condition (1500 ms)				Long cue-probe delay condition (5500 ms)			
Age group	***AX***	***AY***	***BX***	***BY***	***AX***	***AY***	***BX***	***BY***
5-year-olds								
% E	5.6 (3.9)	24.9 (19)	21.8 (20)	13.4 (16)	9.3 (9.4)	37.3 (25)	30.1 (27)	12 (15)
RT	825 (142)	1006 (128)	862 (204)	801 (177)	914 (146)	1098 (190)	818 (203)	902 (202)
Mean *z*-score	−0.18 (0.24)	+0.47 (0.35)	−0.02 (0.46)	−0.27 (0.31)	+0.11 (0.23)	+0.79 (0.38)	−0.16 (0.55)	+0.07 (0.47)
6-year-olds								
% E	3.3 (2.7)	29 (29.8)	6.7 (7.2)	4.2 (7.4)	6.3 (4.4)	36.5 (33.6)	12.8 (11.9)	5.2 (8.2)
RT	687 (154)	893 (153)	670 (227)	663 (183)	762 (154)	984 (89)	693 (207)	710 (178)
Mean *z*-score	−0.17 (0.23)	+0.83 (0.87)	−0.25 (0.56)	−0.28 (0.47)	+0.14 (0.22)	+0.97 (0.43)	−0.18 (0.60)	−0.10 (0.47)

SDs are in parentheses.

rates and on *z*-scores, $F(2,110) = 4.67$, $p < 0.05$, $\eta_p^2 = 0.07$, and $F(2,110) = 4.67$, $p < 0.05$, $\eta_p^2 = 0.07$. They are explored further below.

Does age affect the temporal dynamics of control?

Planned comparisons revealed that younger children produced more errors than older children on BX trials ($M = 25.9\%$, and $M = 9.8\%$, respectively), $F(1,55) = 18.10$, $p < 0.001$, $\eta_p^2 = 0.24$, and on BY trials ($M = 12.6\%$, and $M = 4.7\%$, respectively), $F(1,55) = 7.41$, $p < 0.01$, $\eta_p^2 = 0.11$, whereas error rates between the two age groups did not differ on AY trials, $F(1,55) = 0.07$, $p = 0.79$. Planned comparisons in each age group indicated that the typical proactive pattern observed when considering all participants was observed in the older group only with more errors on AY ($M = 32.8\%$) than on BX trials ($M = 9.8\%$), $F(1,55) = 20.85$, $p < 0.001$, $\eta_p^2 = 0.27$ (see **Figure 2**). In contrast, no significant difference was observed between AY and BX trials ($M = 31\%$ and $M = 25.9\%$, respectively) in 5-year-olds, $F(1,55) = 1.06$, $p = 0.30$. Turning to *z*-scores, planned comparisons revealed that both 5- and 6-year-olds presented longer latencies on AY than on BX trials, $F(1,55) = 37.97$, $p < 0.001$, $\eta_p^2 = 0.40$, and $F(1,55) = 86.64$, $p < 0.001$, $\eta_p^2 = 0.61$, respectively. However, the difference between latencies on AY and BX trials increased from age 5 to 6, $F(1,55) = 5.38$, $p < 0.05$, $\eta_p^2 = 0.08$. The larger difference between AY and BX trials performance was due a difference between age groups latencies on AY trials: on this trial type, 6-year-olds produced slower latencies ($M = 0.90$) than 5-year-olds ($M = 0.63$), $F(1,55) = 6.89$, $p < 0.05$, $\eta_p^2 = 0.11$. Latencies on BX trials ($M = -0.09$ and $M = -0.21$, respectively) and BY trials ($M = -0.10$ and $M = -0.19$, respectively) did not differ between the younger and the older age group, $F(1,55) = 1.23$, $p = 0.27$, and $F(1,55) = 1.80$, $p = 0.18$, respectively.

Considering that the lack of difference between performance on AY and BX trials in the 5-year-old group could not be interpreted, we explored their performance on these trials further in order to investigate whether there might be two subgroups with differing modes of control. We performed a median split based on the critical difference between the error rates observed in these two kinds of trials. It was plausible that none of the subgroups used a reactive mode of control, and that the average difference between AY and BX trials error rates would remain close to zero in both subgroups. Alternatively, the subgroups could differ in their mode of control: one could have performed the task using reactive control, in which case their AY-BX average should be significantly negative, while the other used a proactive mode and thus should have a significantly positive AY-BX average. An ANOVA was run with group (above vs. below the median difference score) as a between-subjects factor and trial type (AY, BX) and delay (1500 ms short vs. 5500 ms long) as within-subjects factors.

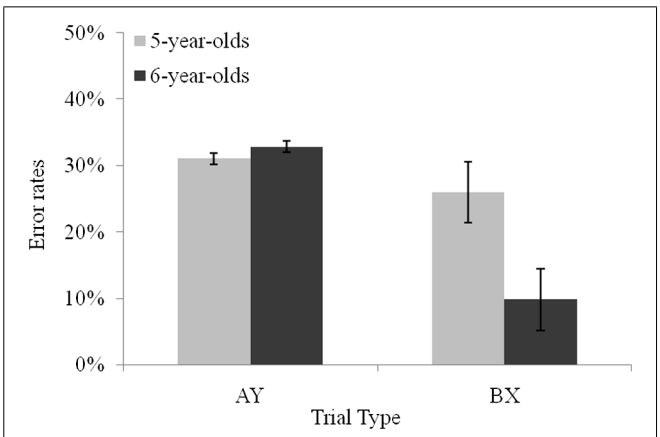


FIGURE 2 | Performance of 5- and 6-year-olds on AY and BX non-target trial types. Error bars reflect SEs of the means.

The analysis revealed a significant interaction between trial type and group, $F(1,27) = 32.29$, $p < 0.001$, $\eta_p^2 = 0.54$. Planned comparisons indicated that on average, participants in the group above the median made more errors on AY trials ($M = 38.8\%$) than on BX trials ($M = 16.3\%$), $F(1,27) = 27.97$, $p < 0.001$, $\eta_p^2 = 0.50$,

which suggests a proactive use of control. The reverse pattern was observed in the group below the median, $F(1,27) = 7.31$, $p < 0.05$, $\eta_p^2 = 0.21$, with more errors in BX trials ($M = 34.9\%$) than of AY trials ($M = 23.7\%$), thereby revealing the use of a reactive form of control (see **Figure 3**)[2]. We also tested whether this contrast between the two subgroups would persist when considering *z*-scores. A new ANOVA was run with group (above vs. below the median difference score) as a between-subjects factor and trial type (AY, BX) and delay (1500 ms short vs. 5500 ms long) as a within-subjects factor. A significant interaction between trial type and group was obtained, $F(1,27) = 5.11$, $p < 0.05$, $\eta_p^2 = 0.15$. Both groups were slower on AY trials than on BX trials, $F(1,27) = 53.82$, $p < 0.001$, $\eta_p^2 = 0.66$, and $F(1,27) = 18.82$, $p < 0.001$, $\eta_p^2 = 0.41$. However, planned comparisons revealed that the difference between latencies on AY and BX trials was larger in the above-median-group than in below-median group, $F(1,27) = 5.11$, $p < 0.05$, $\eta_p^2 = 0.15$. Moreover, in order to gain a better understanding of children's proactive vs. reactive characteristics; we compared children's speed of processing of the two subgroups through latencies on BY trials. This trial is considered as a baseline condition because both cue and probe are associated to non-target responses. Children shown to use reactive control were marginally slower in the more demanding condition (i.e., in the long delay) than children engaging proactive control ($M = 0.22$, and $M = -0.08$, respectively), $t(27) = -1.84$, $p = 0.06$.

In summary, age-related differences were found both on error rates and on *z*-scores. Error rates analyses revealed important inter-individual differences within the 5-year-olds group and altogether these findings shaped a developmental path towards an increasing efficiency of proactive control with age.

Does the cue maintenance delay affect the temporal dynamics of control?

Planned comparisons revealed more errors with the long delay than with the short one on AY trials ($M = 36.9\%$ and $M = 26.9\%$, respectively), $F(1,55) = 9.76$, $p < 0.01$, $\eta_p^2 = 0.15$, and on BX trials ($M = 21.5\%$ and $M = 14.2\%$, respectively), $F(1,55) = 5.28$, $p < 0.05$, $\eta_p^2 = 0.08$, whereas error rates on BY did not differ between the two delays ($M = 8.6\%$ and $M = 8.8\%$, respectively), $F(1,55) = 0.01$, $p = 0.90$. Turning to *z*-scores, planned comparisons showed longer latencies on AY trials with a long delay than with a short delay ($M = 0.88$ and $M = 0.65$, respectively), $F(1,55) = 5.04$, $p < 0.05$, $\eta_p^2 = 0.08$, whereas *z*-scores on BX trials did not differ between the two delay conditions ($M = -0.17$ and $M = -0.14$, respectively), $F(1,55) = 0.12$, $p = 0.72$.

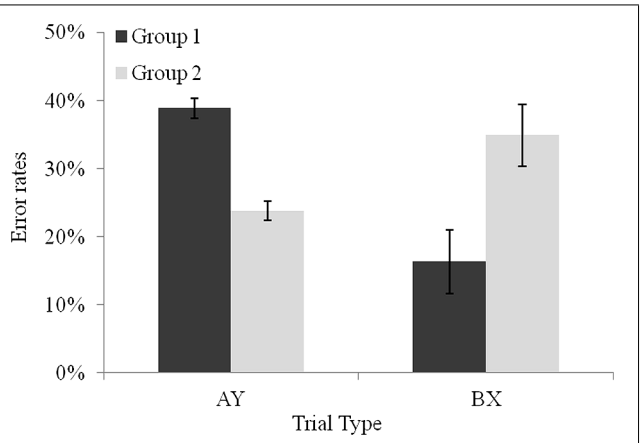


FIGURE 3 | Performance of the two groups of 5-year-olds on AY and BX non-target trial types. Error bars reflect SEs of the means.

THE DEVELOPMENT OF CONTEXT SENSITIVITY (d′ SCORES)

In order to assess the development of children's sensitivity to the preceding context when presented with an X probe, the signal detection index d' was computed (Lorsbach and Reimer, 2008, 2010) corresponding to a ratio between the proportion of correct responses on AX trials (hits) and the proportion of incorrect target responses on BX trials (false alarms). It should be noted that this index does not indicate whether participants use reactive or proactive control to perform the task since false alarms on BX trials can be either due to failures in actively maintaining the B cue, or by a failure to retrieve B cue after the occurrence of X probe. The higher the value of d', the more efficiently the participant used previous goal-related information (A or non-A) to produce a target or a non-target response in response to the X probe. To compare whether 5-year-old children differed from 6-year-olds in their sensitivity to cue information, we ran an ANOVA on d' values with age (5-year-olds vs. 6-year-olds) as a between-subjects variable and delay (1500 ms, short vs. 5500 ms, long) as a within-subjects variable. A main effect of age was observed, $F(1,55) = 23.86$, $p < 0.001$, $\eta_p^2 = 0.18$, with larger d' scores in 6-year-olds than in 5-year-olds ($M = 0.39$ and $M = 0.31$, respectively). A main effect of delay was also found, $F(1,55) = 12.89$, $p < 0.001$, $\eta_p^2 = 0.18$, showing larger d' scores in the short than in the long delay condition ($M = 0.37$ and $M = 0.33$, respectively). However, the interaction between these two variables was not significant, $F(1,55) = 0.11$, $p = 0.73$.

Altogether, results on d' scores revealed an increase in children's sensitivity to cue information between the ages of 5 and 6. In addition, all age groups showed reduced sensitivity to cue information under the long cue–probe delay condition.

DISCUSSION

It is now well established that executive control dramatically develops before the age of 6. Several recent studies converge to suggest that this progress might be sustained by a growing efficiency in activating one's task goal and in maintaining its representation to guide the production of a response. However, the extent to which these changes are supported by a shift in the mode of control used remains under-investigated. The current study aimed

[2]Although we acknowledge that the median-split approach may have maximized the chances of evidencing a positive difference in one subgroup vs. a negative difference in the other, such contrasted patterns of performance were not obtained when conducting the same method in 6-year-olds. The same analysis on error rates as the one run in 5-year-olds revealed a significant interaction between trial type and group, $F(1,26) = 38.96$, $p < 0.001$; $\eta_p^2 = 0.59$. More errors on AY than on BX trials were found in the group above the median, $F(1,26) = 79.00$, $p < 0.001$; $\eta_p^2 = 0.75$, whereas performance between AY and BX trials did not significantly differ in the group below the median ($p = 0.95$). This suggests that the median split in itself does not systematically lead to a conclusion of reactive control.

to (a) explore the temporal dynamics of executive control at the ages of 5 and 6; and (b) study whether manipulating the working memory load influences these dynamics or modulates their efficiency. Our results provide empirical evidence for both qualitative and quantitative changes in the dynamics of control. Importantly, the findings reveal a qualitative shift from reactive to proactive control at the age of 5, as well as graded changes in proactive control from 5- to 6-year-olds. With respect to our second aim, increasing the working memory load did not prevent the active maintenance of goal information; however, it reduced children's sensitivity to the nature of the cue presented earlier. The present results are in accordance with those of previous studies attesting to developmental improvements in activation and maintenance of goals during childhood (Marcovitch et al., 2007, 2010; Chatham et al., 2009; Chevalier and Blaye, 2009; Lorsbach and Reimer, 2010). Further, our findings reveal that the improvement between the ages of 5 and 6 reflects both qualitative and quantitative changes in control. Together, the two groups of children demonstrated the engagement of proactive control, both on error rates and latencies when contrasting their performance on BX and AY trials. We recall that proactive control is reflected by worse performance on AY trials since maintaining cue information is detrimental in this condition due to the high frequency of AX pairs in the task that induces a strong expectation of a target response which then needs to be inhibited when the Y probe is displayed. Whereas this pattern was maintained when considering the older group of children, the picture was less clear-cut in 5-year-olds, who produced similar performance on both types of trials. Further analyses, discussed below, revealed that this mixed picture was probably the consequence of inter-individual differences among this age group. More gradual, quantitative differences were observed between younger and older children. As expected, 6-year-olds appeared more sensitive to cue information in deciding whether or not to produce a target response, corresponding to an increased sensitivity index and less errors on BX trials. They also took longer than 5-year-olds in selecting the non-target response on AY trials. Altogether, these results suggest that context information was better maintained and guided more closely responses in 6-year-olds.

According to the DMC theory, the activation and maintenance of goal representation is underlain by neurobiological mechanisms (lateral PFC and DA system). Proactive control involves a sustained activation of the lateral PFC through a phasic signal of DA, which regulates access of information to enable the active maintenance of task-relevant goal information. In contrast, reactive control is related to a transient activation of PFC because bursts of DA do not occur. During the last decades, behavioral and anatomical studies provided evidence that the PFC and DA system reach maturity during adolescence (Casey et al., 2000; Posner et al., 2012) but dramatically develop during early childhood (Giedd et al., 1999; Rueda et al., 2004, 2005; Moriguchi and Hiraki, 2011). How can children, from at least the age of 6, already use proactive control? It can be argued that the neural substrates underlying proactive control in young children might, at least partially, differ from those activated in adolescents and adults due to their still immature PFC and DA system and/or overall neural activation could be larger than in adults. Alternatively, as suggested by the quantitative indexes of an increase in proactive control efficiency between the ages of 5 and 6, it seems plausible that this form of control is still far from optimal in the older group and hence could be sustained by a still partially immature PFC. Neurophysiological evidence in the field of the development of executive control bear support to each of these hypotheses (see Banich et al., 2013; and Larson et al., 2012; for data compatible with the first and second hypothesis, respectively). Further studies are thus required to investigate the extent to which proactive control in children is subserved by neural substrates similar to adults' proactive control.

A deeper investigation within the 5-year-old group revealed contrasting patterns on error rates with some children already engaging a proactive mode of control to perform the task, and others using a reactive mode. While bearing in mind the limitations of the approach used to set-up the subgroups – which may have reinforced inter-individual differences between the modes of control – this finding suggests that the age of 5 might correspond to a transition in the development of control, at least in situations involving an active maintenance and/or a retrieval of context information. In line with children studies arguing accuracy to be a more sensitive measure than RT (Diamond and Kirkham, 2005; Chevalier and Blaye, 2009), analyses on latencies failed to reveal distinct control modes in the two subgroups. However, these analyses evidenced graded differences in the efficiency of proactive control between the two 5-year-old subgroups that were in a direction consistent with findings on error rates. Although both subgroups took longer to correctly respond to AY than BX trials, this difference was more pronounced – as expected from more efficient users of a proactive mode of control – in the subgroup identified as proactive on the basis of error rates. As proactive control requires maintenance of information during the cue probe delay, while reactive control does not, we considered that reactive patterns could be produced by children less efficient at maintaining information. Research on the development of working memory has established correlations between working memory and speed of processing scores (e.g., Barrouillet et al., 2009; Camos and Barrouillet, 2011). Indeed, the two subgroups contrasted here revealed marginal differences in terms of speed of processing. As expected, children shown to use reactive control were slower in the more demanding condition (i.e., long delay). Although further investigation of their working memory capacities would be necessary, this finding offers a convergent pattern with the error rate analysis. We will discuss further the relations between mode of control and working memory when considering the effect of the delay between cue and probe. We now examine recent results published independently while this study was run that suggest that a shift between reactive and proactive control might occur one year later that is, at 6 years of age.

Blackwell and Munakata (2013) suggested that the dynamics of control can be evidenced by considering children's performance in a three dimensional version of the DCCS (3-DCCS). In this task, participants have to sort tridimensional stimuli. This leads to three blocks of trials, each block corresponding to one type of sort imposed by the experimenter's instructions (i.e., sorting

first by shape, then by color, then by size). The authors' reasoning is that children who succeed to switch from one block to another use a proactive control because they achieve to maintain the relevant sorting rule which is given only once at the beginning of each block in a highly interfering context due to the two other rules. By contrast, perseveration would reveal a difficulty of reactivating the correct sorting rule in this highly conflictive context, authorizing to consider perseverators as engaging a reactive control. It might be argued that the age difference in the transition from reactive to proactive control between this study and the current one is due to the index used: switching between tasks through post-switch accuracy vs. performance on AY and BX trials. However, on the one hand the AX-CPT is the most characteristic task to assess the dynamics of control, and on the other hand, Blackwell and Munakata's (2013) findings revealed that the *a priori* categorization of switchers as proactive and perseverators as reactive was corroborated by their performance on a delayed matching task. Hence, a new question must be raised: could the differences between the two tasks used to contrast the two modes of control account for the one year difference to observe a shift across the two tasks? We contend that the 3-DCCS is more demanding in terms of active maintenance since the tridimensional stimuli trigger not only the currently relevant rules but also the two irrelevant ones. By contrast the AX-CPT makes proactive control easier to engage since participants do not encounter any stimuli during the cue-maintenance delay.

Given the limited working-memory capacity in young children, we hypothesized that increasing the working-memory load through lengthening the delay of cue maintenance would increase the working memory load and hence, would decrease children's efficiency at using the cue to guide their response to the probe, thereby inducing a shift from a proactive to a reactive mode of control. The results did not support this hypothesis since no reversal of the pattern of control was observed. This could suggest that active maintenance of goal-related information from the cue is not the most critical determinant of the mode of control engaged, at least in the age groups studied. Instead, it could be more crucial to retrieve an explicit representation of the goal when seeing the cue. Recent research by Chevalier and Blaye (Chevalier and Blaye, 2009; Blaye and Chevalier, 2011) has pointed to the role of task–cue translation into goals in preschoolers' performance on a flexibility task. By contrasting different types of task–cues that varied in their degree of transparency (i.e., the degree of association between cue and task goal), the authors demonstrated that preschoolers had specific difficulties to retrieve a representation of what they had to do next when arbitrary cues were used even though they were able to recall the meaning of the cues. Cues used in the AX-CPT are arbitrary; the pairs presented as target pairs (AX) or non-target pairs (AY, BX, and BY) are all arbitrarily composed and the expected response has no relation with the animals either (i.e., pressing a green or red button). Hence, it would be worth comparing the mode of control engaged in different versions of the AX-CPT by a same sample of preschoolers depending on whether the cues–probes–responses associations would arbitrary or meaningful. Such a meaningful version has been used by Chatham et al. (2009) but on different age groups. Interestingly, Lorsbach and Reimer (2010) interpreted 8-year-olds' weaker proactive control, compared to older children, as arising from difficulties to transform the cue into a complete representation of the goal.

A more parsimonious interpretation of the lack of shift from one mode of control to another when contrasting the two cue–probe delays could be that, the two delays are either too much or not sufficiently demanding in terms of maintenance. The overall proactive control observed in the two age groups does not support the hypothesis of two delays that would be too demanding; however, this might be at least partly the case for the 5-year-old subgroup that was found to use a reactive mode of control in both delay conditions. Alternatively, one may assume that increasing the cue–probe delay without any additional information to process in the meanwhile is not sufficiently demanding to induce qualitative changes in control. It could be worth testing the effect of another form of WM load manipulation, namely varying the demand of a concurrent processing task during the cue-probe delay. Nevertheless, the absence of a shift in the dynamics of control when lengthening the cue probe delay does not mean a lack of impact of this manipulation. More graded measures revealed quantitative changes suggesting that manipulating the delay does affect the working memory load. Children's efficiency in using the cue information to guide their response to the probe appeared to be lowered with longer delay. Hence, when goal-related information has to be actively maintained, preschool age children can encounter difficulties to use it, without demonstrating the use of a pure reactive mode of control.

In sum, the current study aimed to investigate some of the quantitative and qualitative changes in activation and maintenance of goal representation between 5 and 6 years of age that might sustain the development of executive control. Although two recent theoretical papers (Braver, 2012; Munakata et al., 2012) offered the hypothesis that the development of executive control could correspond to a shift from reactive to proactive control during childhood, empirical validation of this hypothesis remains scarce. The current study proposes a children-adapted version of the AX-CPT and suggests that such a shift might occur at 5 years of age. This new finding is somewhat at odds with the results obtained by Blackwell and Munakata (2013). These authors observed this transition one year later using a different task originally designed to assess flexibility. This décalage raises the question of the extent to which this reversal in the temporal dynamics of control depends on the task demand in terms of active maintenance of goal information. Future investigation of this question should lead to a more complex picture of the development of executive control than the probably too simplistic view suggesting that these two modes of control correspond to two developmental stages.

AUTHOR CONTRIBUTIONS

Joanna Lucenet and Agnès Blaye designed the Experiment. Data collection was carried out by Joanna Lucenet. Joanna Lucenet drafted the manuscript and Agnès Blaye provided critical revisions. Joanna Lucenet and Agnès Blaye have all approved the final version of the manuscript and agree to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

The present research was funded by the Agence Nationale de la Recherche (ANR) through grants to Agnès Blaye (ANR-07-FRAL-015 and ANR-ANAFONEX-BLAN-1908-02). Special thanks to Maria Ktori and Sebastiaan Mathôt for helpful comments.

REFERENCES

Baddeley, A., Chincotta, D., and Adlam, A. (2001). Working memory and the control of action: evidence from task switching. *J. Exp. Psychol. Gen.* 130, 641–657. doi: 10.1037/0096-3445.130.4.641

Banich, M. T., De La Vega, A., Andrews-Hanna, J. R., Seghete, K. M., Du, Y. P., and Claus, E. D. (2013). Developmental trends and individual differences in brain systems involved in intertemporal choice during adolescence. *Psychol. Addict. Behav.* 27, 416–430. doi: 10.1037/a0031991

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., and Camos, V. (2009). Working memory span development: a time-based resource-sharing model account. *Dev. Psychol.* 45, 477–490. doi: 10.1037/a0014615

Blackwell, K. A., and Munakata, Y. (2013). Costs and benefits linked to developments in cognitive control. *Dev. Sci.* 17, 203–211. doi: 10.1111/desc.12113

Blair, C., and Razza, R. P. (2007).Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Blaye, A., and Chevalier, N. (2011). The role of goal representation in preschoolers' flexibility and inhibition. *J. Exp. Child Psychol.* 108, 469–483. doi: 10.1016/j.jecp.2010.09.006

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci.* 16, 106–113. doi: 10.1016/j.tics.2011.12.010

Braver, T. S., and Barch, D. A. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neurosci. Biobehav. Rev.* 26, 809–817. doi: 10.1016/S0149-7634(02)00067-2

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., et al. (2001). Context processing in older adults: evidence for a theory relating cognitive control to neurobiology in healthy aging. *J. Exp. Psychol. Gen.* 130, 746–763. doi: 10.1037/0096-3445.130.4.746

Braver, T. S., Gray, J. R., and Burgess, G. C. (2007). "Explaining the many varieties of working memory variation: dual mechanisms of cognitive control," in *Variation in Working Memory*, eds A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (Oxford: Oxford University Press), 76–106.

Braver, T. S., Satpute, A. B., Rush, B. K., Racine, C. A., and Barch, D. M. (2005).Context processing and context maintenance in healthy aging and early stage dementia of the Alzheimer's type. *Psychol. Aging* 20, 33–46. doi: 10.1037/0882-7974.20.1.33

Camos, V., and Barrouillet, P. (2011). Developmental change in working memory strategies: from passive maintenance to active refreshing. *Dev. Psychol.* 47, 898–904. doi: 10.1037/a0023193

Cannard, C., Blaye, A., Scheuner, N., and Bonthoux, F. (2005). Picture naming in 3- to 8-year-old French children: methodological considerations for name agreement. *Behav. Res. Methods* 37, 417–425. doi: 10.3758/BF03192710

Carlson, S. A. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M., Mandell, D. J., and Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Dev. Psychol.* 40, 1105–1122. doi: 10.1037/0012-1649.40.6.1105

Carlson, S. M., and Moses, L. J. (2001).Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333

Carlson, S. M., and Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cogn. Dev.* 22, 489–510. doi: 10.1016/j.cogdev.2007.08.002

Casey, B. J., Giedd, J. N., and Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biol. Psychol.* 54, 241–257. doi: 10.1016/S0301-0511(00)00058-2

Chalard, M., Bonin, P., Meot, A., Boyer, B., and Fayol, M. (2003). Objective age-of-acquisition (AoA) norms for a set of 230 object names in French: relationships with psycholinguistic variables, the English data from Morrison et al. (1997), and naming latencies. *Eur. J. Cogn. Psychol.* 15, 209–245. doi: 10.1080/09541440244000076

Chatham, C. H., Frank, M. J., and Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5529–5533. doi: 10.1073/pnas.0810002106

Chevalier, N., and Blaye, A. (2009). Setting goals to switch between tasks: effect of cue transparency on children's cognitive flexibility. *Dev. Psychol.* 45, 782–797. doi: 10.1037/a0015409

Chevalier, N., Blaye, A., Dufau, S., and Lucenet, J. (2010). What Visual Information Do Children and Adults Consider While Switching Between Tasks? Eye-Tracking Investigation of Cognitive Flexibility Development. *Dev. Psychol.* 46, 955–972. doi: 10.1037/a0019674

Chevalier, N., Huber, K. L., Wiebe, S. A., and Espy, K. A. (2013). Qualitative change in executive control during childhood and adulthood. *Cognition* 128, 1–12. doi: 10.1016/j.cognition.2013.02.012

Chevalier, N., Sheffield, T. D., Nelson, J. M., Clark, C. A. C., Wiebe, S. A., and Espy, K. A. (2012). Underpinnings of the costs of flexibility in preschool children: the roles of inhibition and working memory. *Dev. Neuropsychol.* 37, 99–118. doi: 10.1080/87565641.2011.632458

Chevalier, N., Wiebe, S. A., Huber, K. L., and Espy, K. A. (2011). Switch detection in preschoolers' cognitive flexibility. *J. Exp. Child Psychol.* 109, 353–370. doi: 10.1016/j.jecp.2011.01.006

Clark, C. A. C., Sheffield, T. D., Wiebe, S. A., and Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Dev.* 84, 662–677. doi: 10.1111/j.1467-8624.2012.01854.x

Dauvier, B., Chevalier, N., and Blaye, A. (2012). Using finite mixture of GLMs to explore variability in children's flexibility in a task-switching paradigm. *Cogn. Dev.* 27, 440–454. doi: 10.1016/j.cogdev.2012.07.004

Deak, G. O. (2003). The development of cognitive flexibility and language abilities. *Adv. Child Dev. Behav.* 31, 271–327. doi: 10.1016/S0065-2407(03)31007-9

Diamond, A., and Kirkham, N. (2005). Not quite as grown-up as we like to think – parallels between cognition in childhood and adulthood. *Psychol. Sci.* 16, 291–297. doi: 10.1111/j.0956-7976.2005.01530.x

Duncan, J., Emslie, H., Williams, P., Johnson, R., and Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn. Psychol.* 30, 257–303. doi: 10.1006/cogp.1996.0008

Eisenberg, N., and Sulik, M. J. (2012). Emotion-related self-regulation in children. *Teach. Psychol.* 39, 77–83. doi: 10.1177/0098628311430172

Emerson, M. J., and Miyake, A. (2003). The role of inner speech in task switching: a dual-task investigation. *J. Mem. Lang.* 48, 148–168. doi: 10.1016/S0749-596X(02)00511-9

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *J. Exp. Psychol. Gen.* 137, 201–225. doi: 10.1037/0096-3445.137.2.201

Gathercole, S. E., Pickering, S. J., Ambridge, B., and Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* 40, 177–190. doi: 10.1037/0012-1649.40.2.177

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nat. Neurosci.* 2, 861–863. doi: 10.1038/13158

Gruber, O., and Goschke, T. (2004). Executive control emerging from dynamic interactions between brain systems mediating language, working memory and attentional processes. *Acta Psychol.* 115, 105–121. doi: 10.1016/j.actpsy.2003.12.003

Karbach, J., and Kray, J. (2007). Developmental changes in switching between mental task sets: the influence of verbal labeling in childhood. *J. Cogn. Dev.* 8, 205–236. doi: 10.1080/15248370701202430

Larson, M. J., Clawson, A., Clayson, P. E., and South, M. (2012). Cognitive control and conflict adaptation similarities in children and adults. *Dev. Neuropsychol.* 37, 343–357. doi: 10.1080/87565641.2011.650337

Lehto, J. E., Juujarvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Lorsbach, T. C., and Reimer, J. F. (2008). Context processing and cognitive control in children and young adults. *J. Genet. Psychol.* 169, 34–50. doi: 10.3200/GNTP.169.1.34-50

Lorsbach, T. C., and Reimer, J. F. (2010). Developmental differences in cognitive control: goal representation and maintenance during a continuous performance task. *J. Cogn. Dev.* 11, 185–216. doi: 10.1080/15248371003699936

Marcovitch, S., Boseovski, J. J., and Knapp, R. J. (2007). Use it or lose it: examining preschoolers' difficulty in maintaining and executing a goal. *Dev. Sci.* 10, 559–564. doi: 10.1111/j.1467-7687.2007.00611.x

Marcovitch, S., Boseovski, J. J., Knapp, R. J., and Kane, M. J. (2010). Goal neglect and working memory capacity in 4- to 6-year-old children. *Child Dev.* 81, 1687–1695. doi: 10.1111/j.1467-8624.2010.01503.x

Miller, S. E., and Marcovitch, S. (2012). How theory of mind and executive function co-develop. *Rev. Philos. Psychol.* 3, 597–625. doi: 10.1007/s13164-012-0117-0

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moriguchi, Y., and Hiraki, K. (2011). Longitudinal development of prefrontal function during early childhood. *Dev. Cogn. Neurosci.* 1, 153–162. doi: 10.1016/j.dcn.2010.12.004

Morton, J. B., and Munakata, Y. (2002). Active versus latent representations: a neural network model of perseveration, dissociation, and decalage. *Dev. Psychobiol.* 40, 255–265. doi: 10.1002/dev.10033

Munakata, Y., Snyder, H. R., and Chatham, C. H. (2012). Developing cognitive control: three key transitions. *Curr. Dir. Psychol. Sci.* 21, 71–77. doi: 10.1177/0963721412436807

Paxton, J. L., Barch, D. M., Racine, C. A., and Braver, T. S. (2008). Cognitive control, goal maintenance, and prefrontal function in healthy aging. *Cereb. Cortex* 18, 1010–1028. doi: 10.1093/cercor/bhm135

Posner, M. I., Rothbart, M. K., Sheese, B. E., and Voelker, P. (2012). Control networks and neuromodulators of early development. *Dev. Psychol.* 48, 827–835. doi: 10.1037/a0025530

Redick, T. S. (2014). Cognitive control in context: working memory capacity and proactive control. *Acta Psychol.* 145, 1–9. doi: 10.1016/j.actpsy.2013.0.010

Rubinstein, J. S., Meyer, D. E., and Evans, J. E. (2001). Executive control of cognitive processes in task switching. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 763–797. doi: 10.1037/0096-1523.27.4.763

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., and Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14931–14936. doi: 10.1073/pnas.0506897102

Towse, J. N., Lewis, C., and Knowles, M. (2007). When knowledge is not enough: the phenomenon of goal neglect in preschool children. *J. Exp. Child Psychol.* 96, 320–332. doi: 10.1016/j.jecp.2006.12.007

Wright, I., Waterman, M., Prescott, H., and Murdoch-Eaton, D. (2003). A new Stroop-like measure of inhibitory function development: typical developmental trends. *J. Child Psychol. Psychiatry* 44, 561–575. doi: 10.1111/1469-7610.00145

Zelazo, P. D., Frye, D., and Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cogn. Dev.* 11, 37–63. doi: 10.1016/S0885-2014(96)90027-1

Zelazo, P. D., Müller, U., Frye, D., and Marcovitch, S. (2003). The development of executive function in early childhood: VI. The development of executive function: cognitive complexity and control–revised. *Monogr. Soc. Res. Child Dev.* 68, 93–119. doi: 10.1111/j.0037-976X.2003.00266.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2014; accepted: 12 July 2014; published online: 29 July 2014.
*Citation: Lucenet J and Blaye A (2014) Age-related changes in the temporal dynamics of executive control: a study in 5- and 6-year-old children. Front. Psychol. **5**:831. doi: 10.3389/fpsyg.2014.00831*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Lucenet and Blaye. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 17 February 2014
doi: 10.3389/fpsyg.2014.00107

Gaining control: changing relations between executive control and processing speed and their relevance for mathematics achievement over course of the preschool period

Caron A. C. Clark[1]*, Jennifer Mize Nelson[2], John Garza[2], Tiffany D. Sheffield[2], Sandra A. Wiebe[3] and Kimberly Andrews Espy[4]

[1] *Department of Psychology and Prevention Science Institute, University of Oregon, Eugene, OR, USA*
[2] *Developmental Cognitive Neuroscience Laboratory, Department of Psychology and Office of Research, University of Nebraska-Lincoln, Lincoln, NE, USA*
[3] *Department of Psychology, University of Alberta, Edmonton, AB, Canada*
[4] *Department of Psychology, University of Oregon, Eugene, OR, USA*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Ruth Ford, Anglia Ruskin University, UK
Daniel Berry, University of Illinois, Urbana-Champaign, USA

****Correspondence:***
Caron A. C. Clark, Department of Psychology and Prevention Science Institute, University of Oregon, Rm. 428, Lewis Integrative Science Building, Eugene, OR 97403, USA
e-mail: carrie4@uoregon.edu

Early executive control (EC) predicts a range of academic outcomes and shows particularly strong associations with children's mathematics achievement. Nonetheless, a major challenge for EC research lies in distinguishing EC from related cognitive constructs that also are linked to achievement outcomes. Developmental cascade models suggest that children's information processing speed is a driving mechanism in cognitive development that supports gains in working memory, inhibitory control and associated cognitive abilities. Accordingly, individual differences in early executive task performance and their relation to mathematics may reflect, at least in part, underlying variation in children's processing speed. The aims of this study were to: (1) examine the degree of overlap between EC and processing speed at different preschool age points; and (2) determine whether EC uniquely predicts children's mathematics achievement after accounting for individual differences in processing speed. As part of a longitudinal, cohort-sequential study, 388 children (50% boys; 44% from low income households) completed the same battery of EC tasks at ages 3, 3.75, 4.5, and 5.25 years. Several of the tasks incorporated baseline speeded naming conditions with minimal EC demands. Multidimensional latent models were used to isolate the variance in executive task performance that did not overlap with baseline processing speed, covarying for child language proficiency. Models for separate age points showed that, while EC did not form a coherent latent factor independent of processing speed at age 3 years, it did emerge as a distinct factor by age 5.25. Although EC at age 3 showed no distinct relation with mathematics achievement independent of processing speed, EC at ages 3.75, 4.5, and 5.25 showed independent, prospective links with mathematics achievement. Findings suggest that EC and processing speed are tightly intertwined in early childhood. As EC becomes progressively decoupled from processing speed with age, it begins to take on unique, discriminative importance for children's mathematics achievement.

Keywords: executive function, preschool, academic achievement, processing speed, mathematics

INTRODUCTION

Measures of executive control (EC) have gained increasing popularity in developmental science, due in part to their strong ability to predict children's school readiness and academic achievement. For instance, children's performance on executive tasks in preschool correlates with their mathematics achievement well into elementary school (Bull et al., 2008; Clark et al., 2010; LeFevre et al., 2013). So compelling are these predictive relations that they have spurred the development of intervention programs aimed at boosting children's EC prior to school entry (e.g., Diamond et al., 2007; Bierman et al., 2008). Unfortunately, this powerful evidence for the predictive utility of executive tasks contrasts with a relatively limited understanding of the fundamental nature and development of EC as a latent construct. By definition, EC recruits and orchestrates other cognitive processes to facilitate goal-directed behavior. Measures designed to assess EC therefore are multidimensional and draw on an array of basic information processing skills, making it difficult to isolate the precise role of EC in manifest performance (Rabitt, 1997; Miyake et al., 2000; Chan et al., 2008). This conflation of EC with general information processing may be especially problematic in early childhood, when executive tasks necessarily require varied stimuli and response demands and a high degree of verbal scaffolding to promote engagement. To clearly specify the unique implications of early EC for children's academic achievement, we first need to understand how EC intersects with and diverges

from basic processing abilities that also shape children's academic trajectories.

In global theories of cognitive development, processing speed is conceptualized as a central mental capacity that drives changes in higher-order cognition (Hale, 1990; Kail and Salthouse, 1994). Growth in processing speed, as assessed using simple measures of reaction time, follows a predictable, exponential pattern, independent of individual task stimuli or response requirements (Kail, 1991a,b). These age-related gains in processing speed are thought to facilitate general cognitive efficiency in two ways: (1) a greater amount of information can be absorbed within a given time frame and (2) with less time for information to decay, a larger number of neural networks can be co-activated, increasing the capacity to carry out simultaneous operations and represent information from multiple standpoints (Salthouse, 1996). Age-related changes in global processing speed therefore are argued to trigger cascading effects on higher-order systems like EC by constraining or enhancing the efficiency with which information can be processed in a domain-general manner (Kail and Salthouse, 1994; Fry and Hale, 1996).

Findings from several studies support this developmental cascade hypothesis. In literature on aging, processing speed has been found to explain an average of 75% of the variance in elderly adults' performance decline across a variety of complex cognitive tasks (Salthouse, 1996). Processing speed also accounts for between 70 and 90% of the age-related variance in fluid intelligence quotients in children and adults (Kail and Salthouse, 1994; Grudnik and Kranzler, 2001). More specific to EC, measures of processing speed have been shown to fully mediate the relation of age to inhibitory control task performance (Kail, 2002; McAuley and White, 2011) and to partially mediate the relation of age to working memory in middle childhood (Fry and Hale, 2000; McAuley and White, 2011). A seminal study by Case et al. (1982) showed that when experimental manipulations were used to equate adults and 6-year old children in their speed of information processing, their average working memory spans were equivalent. Likewise, using a latent modeling approach, where executive tasks were loaded simultaneously on EC and processing speed factors, Span et al. (2004) found that adults and school-aged children differed only in their mean processing speed and not in latent EC. Recently, Rose et al. (2011) used structural equation modeling to test the cascade model in children born preterm and full term. Consistent with a cascade effect, processing speed mediated the relation between preterm birth and impairments in EC, which in turn were associated with lower reading and mathematics achievement.

Collectively, the above studies support the idea that limitations in processing speed may constrain an individual's ability to perform more complex cognitive tasks, including the inhibition, maintenance and shifting operations attributed to EC. In fact, one conceptual model of EC includes processing speed as a key component of the executive system (Anderson, 2008; Anderson and Reidy, 2012). To date, however, no studies have examined the degree of overlap between processing speed and EC in very young children, despite the fact that increases in both processing speed and executive task performance are especially rapid during early childhood (Kail, 1991a; Wiebe et al., 2012). Given that processing speed is so ubiquitously involved in cognitive task performance, it is possible that a large proportion of the variance in young children's early executive task performance, as well as the relation of executive performance to academic achievement, may be explained by individual differences in processing speed. Addressing this question of overlap is important from a psychometric standpoint, as it challenges the very notion of EC as an independent dimension of cognition, suggesting that executive measures may not capture anything distinct from what is captured by general measures of processing speed (Salthouse et al., 2003; Fournier-Vicente et al., 2008). From a broader theoretical perspective, understanding the early relations between processing speed and EC may also yield important insights into the nature of EC development. It is conceivable, for instance, that rapid changes in myelination, synaptogenesis and connectivity during early childhood might promote system-wide changes in processing speed that facilitate executive performance in a bottom-up manner. On the other hand, temporally specific changes in frontal neural systems may promote relatively discrete age-related advancements in EC independent of gains in processing speed (Span et al., 2004). Clearly, these different developmental mechanisms would also suggest either more general or more specific strategies for early intervention.

One methodological approach that has proven powerful in understanding the underlying nature of EC at different stages of development is confirmatory factor analysis (CFA). The primary advantage of CFA is that it isolates the shared variance from several cognitive tasks that are selected a-priori to measure a given construct, thereby enhancing measurement precision and reducing error. Using CFA of executive tasks administered to school-aged children and adults, studies generally have identified 2–3 distinct but correlated factors that are conceptualized as separate components of EC and typically are labeled inhibitory control, working memory/updating and cognitive flexibility (Miyake et al., 2000; Huizinga et al., 2006; Friedman et al., 2008; Lee et al., 2013). A surprising and replicated finding from CFA studies in preschool-aged children has been the lack of differentiation of EC into distinct components (Wiebe et al., 2008, 2011; Hughes et al., 2010; Willoughby et al., 2010; Fuhs and Day, 2011). Specifically, these studies show that the overlapping variance in preschoolers' executive task performance is most parsimoniously modeled as a single, unitary factor. Collectively, these studies hint at potential changes in the underlying structure of EC over the course of childhood, although constraints on the number and types of executive tasks that can feasibly be administered to young children make it difficult to draw comparisons across different age groups. More importantly, a major limitation of any factor analytic approach is that it is not clear whether the common variance extracted from multiple tasks only reflects the construct of interest. Given that all measures of EC also tap other "bottom-up" processes and that global processing speed is thought to support performance across all higher-order cognitive tasks, it is likely that at least part of the overlap in an individual's performance on different executive tasks that is captured by his or her factor score can be attributed to the general speed with which he or she processes information. Accordingly, the first aim of this study was to use more sophisticated CFA models to parse the relative contributions of EC

and processing speed to young children's executive performance. Manifest executive performance was assumed to reflect a combination of EC, processing speed and other individual differences, as well as task-specific error variance. Each executive task was loaded simultaneously onto an EC and a processing speed factor to capture relative demand on each of these constructs. Language proficiency also was statistically controlled for, given the recognized importance of language for EC development (Wolfe and Bell, 2007; Hughes et al., 2010). As argued by Salthouse et al. (2003), this type of model provides a stringent test of the divergent validity of EC because it directly pits the EC demands against the processing demands of the tasks.

We were particularly interested in whether the contributions of EC and processing speed to children's executive task performance might change over the preschool period. The rationale for this aim stems in part from our longitudinal findings on the structure of EC. At a broad, configural level, the shared variance from a repeatedly administered battery of executive tasks is best modeled as a unitary EC construct regardless of assessment point. At a more nuanced level, this EC factor does not show longitudinal metric or scalar invariance; there are changes in the way that executive tasks relate to the EC construct and in the degree of measurement error over time (Nelson et al., 2014). Cascade models suggest that growth in processing speed frees cognitive resources that then can be devoted to higher-order EC (Case et al., 1982). It is plausible, then, that relative contribution of EC capacities to executive task performance might gradually increase with age-related gains in processing speed. To examine this issue, multidimensional measurement models were fit at different preschool age points and metric invariance tests were performed to describe temporal changes in the EC and processing speed factor loadings.

The final study aim was to determine whether the processing speed demands of executive tasks might drive their relation to mathematics achievement. Strong associations between early EC and mathematics are conceptually appealing because mathematics often involves simultaneous processing and differential allocation of attention—e.g., remember the number of digits counted on one hand while you count the remaining fingers on the other. There is also substantial evidence, however, that children with poorer mathematics achievement generally are slower to process information (Bull and Johnston, 1997; Geary et al., 2012). In studies where covariate approaches have been used to isolate the contributions of EC and processing speed to mathematics achievement, executive measures have sometimes predicted mathematics achievement over and above measures of processing speed (Geary, 2011; Clark et al., 2013). Unfortunately, a covariate approach does not optimally capture the intersecting EC and processing speed demands of the executive tasks themselves. For instance, if executive task performance actually is confounded by underlying variation in processing speed, then two measures of the same construct essentially are competing in the covariate model. Using a CFA approach, van der Sluis et al. (2007) showed that the working memory updating component, but not the cognitive flexibility component of EC, was significantly associated with arithmetic achievement in school aged children after the non-executive demands of EC tasks also were modeled. Notably though, the proportion of arithmetic variance accounted for by the working memory factor was small (2.6%) relative to the proportion accounted for by the non-executive demands of EC tasks (30%), suggesting that the strong relations generally observed between executive task performance and mathematics achievement may largely be driven by the non-executive, baseline processing demands of the executive tasks. Here, we used a similar modeling approach with data from different preschool age points to determine the extent to which the processing speed and EC demands of executive tasks contributed to mathematics achievement, covarying also for language proficiency, over the course of the preschool period.

METHODS

The study included 388 preschoolers (193 boys, 195 girls; 286 Caucasian, 31 Hispanic, 20 African American, 1 Asian, 50 multiracial) drawn from two Midwestern sites, a semi-rural area and a small city. A cohort-sequential design was used to control for practice effects associated with repeated testing; the majority of children ($n = 228$) were enrolled at age 3 years, with smaller numbers enrolled at 3.75 years ($n = 57$), 4.5 years ($n = 55$), and 5.25 years ($n = 48$) respectively. Retention rates for the earlier-recruited cohorts were high (90–100%). Children with developmental impairments (e.g., language delays, Autism) and families whose first language was not English were excluded from recruitment during a preliminary screening call. Families with lower SES were oversampled for greater diversity so that 44.1% of the study families were eligible for public medical assistance or free school lunch or had income levels below Health and Human Services poverty guidelines. Mean length of maternal education at study entry was 14.97 ($SD = 2.37$) years.

PROCEDURE

All procedures were approved by a university institutional review board. At the initial recruitment, researchers visited families' homes to obtain written, informed consent, to observe each child's home environment and to complete the Woodcock—Johnson III Brief Intellectual Ability Assessment (BIA; Woodcock et al., 2001a) with the child. Within a narrow 2-week window, children then visited a university-based laboratory to complete a battery of executive tasks, administered by a trained research technician. These laboratory visits were repeated every 9 months until the child was 5.25 years old. During visits, the child's primary caregiver was interviewed regarding the child's health and family background and also completed several questionnaires related to the child's wellbeing and behavior. At all assessment points, children were administered alternating forms of the Test of Early Mathematics Ability −3 (Ginsburg and Baroody, 2003). Additionally, the Applied Problems subtest from the Woodcock—Johnson III Ability Battery (Woodcock et al., 2001b) was administered at ages 3.75, 4.5, and 5.25 years. At study exit, children were re-administered the BIA.

MEASURES

Executive control and processing speed

A broad array of measures, differing in content and response demands, was chosen to assess putative components of EC, including working memory, inhibitory control, and cognitive

flexibility. A number of these executive tasks also comprised a baseline component or condition, where children were required simply to respond to colors or shapes as quickly as possible and demands on EC theoretically were minimal. Performance on many of the tasks was coded in Noldus Observer by trained undergraduate research assistants, who were blind to study hypotheses. Inter-rater reliability was computed based on 20% of the videos that were randomly selected for independent scoring or cross-coding by another research assistant.

Nine Boxes (adapted from Diamond et al., 1997) was selected to assess working memory. This self-ordered pointing-type task required children to search for hidden figurines in nine boxes with varying colors and lid shapes. During a 15 s delay between selections, the boxes were scrambled behind a screen. The most efficient search strategy entailed selecting only boxes that had not previously been selected. A maximum of 20 trials were administered, the task otherwise ceasing once all of the figurines had been retrieved or once the child had made 5 consecutive errors. Inter-rater reliability was 100%. The single dependent variable for this task was the child's maximum run of consecutive correct responses.

Delayed Alternation (Goldman et al., 1971; Espy, 1999) is a working memory task requiring the child to retrieve a food reward from one of two testing wells covered with neutrally-colored cups. When a child made a correct response, the reward was switched to the opposite well. Between trials, there was a 10 s delay, where the researcher verbally distracted the child while she hid the reward out of view. Three training trials were administered, followed by up to 16 test trials. The task was discontinued after 9 correct responses and the child was given credit for the remaining trials. Inter-rater reliability was 100%. The dependent variable for the task was the maximum length of consecutive incorrect responses subtracted from the maximum length of consecutive correct responses.

Nebraska Barnyard (adapted from Hughes et al., 1998) is a working memory span-type task requiring the child to remember increasing sequences of animal names. The task was programmed in Perl (Active State Software, Vancouver, BC, Canada) and administered on a touch-screen computer. During an initial training phase, children were presented with 9 colored buttons arranged in a grid-like pattern on the computer screen. Each button included a picture of an animal (e.g., green with a frog, pink with a pig) and emitted the sound the corresponding animal sound when pressed. Children were encouraged to memorize each animal's location. Thereafter, the pictures of the animals were removed, leaving only the colored buttons. Children were asked to push the buttons corresponding to progressively increasing sequences of animal names read by the examiner. Up to three trials were administered for each sequence level and children were given automatic credit for the third trial if they correctly completed the first two trials. The task ceased when the child was unable to repeat all three sequences of animal names at a given sequence length. Coding was completed in Noldus; inter-rater reliability was 96%. The dependent variable for this task was the total number of correct trials -1/3rd of a point was added for each correct one-animal sequence.

Big-Little Stroop (Kochanska et al., 2000) assessed processing speed and proactive inhibition and required children to name smaller shapes embedded within a larger shape. The task was administered in EPrime (Psychology Software Tools, Pittsburgh, PA, USA), with black and white line drawings used as stimuli. Of the 24 trials administered, 50% were conflict trials, where the embedded shapes were different from the larger shape and 50% were non-conflict trials, where the embedded shapes matched the larger shape. Prior to the onset of the test stimulus, a brief (750 ms) priming stimulus of the larger shape was presented. Inter-rater reliability was 90% for response times and 99% for accuracy, both of which were coded in Noldus. Dependent variables from this task included mean response times for correct non-conflict trials and mean accuracy for conflict trials.

A *Go/No-Go* task (adapted from Simpson and Riggs, 2006) provided a measure of response inhibition. During this task, children were instructed to press a button when a picture of a fish appeared on the computer screen (75%), but to refrain from pressing the button when a picture of a shark was presented (25% of trials). After each trial, children were shown a net, which appeared broken simultaneous with a buzzing sound if the child made an error of commission. Stimuli were presented in Eprime for 1500 ms, with an inter-stimulus interval of 100 ms. The dependent variable was dPrime (d'; the standardized ratio of hits to misses).

The *Modified Snack Delay* task (adapted from Kochanska et al., 1996; Korkman et al., 1998) was used to assess motor inhibition. Children were instructed to maintain a still posture and remain completely silent with their hands positioned on a mat until the researcher rang a bell after 240 s. A handful of M & M candies was positioned under a transparent glass in front of the child. At specific intervals during the delay, the researcher implemented a scripted set of distracters designed to break the child's pose (e.g., coughing, dropping a pencil, leaving the room for 1.5 min to fetch more candy). Inter-rater reliability was 90%. A hand movement score was used as the dependent variable; children were allocated a point for each epoch with no hand movement, half a point for epochs with some hand movement and 0 points for lots of hand movement. If the child ate the candy, the movement score was calculated based on the epochs completed prior to that point.

A computerized version of the *Shape School* (Espy, 1997) task provided measures of baseline processing speed, response inhibition and cognitive flexibility. Children were presented with cartoon stimuli that varied on the dimensions of color (red, blue), shape (circle, square), emotion (happy, sad), and cue (wearing a hat, not wearing a hat). For the first, baseline task condition (12 trials), children were instructed to name the colors of the characters as quickly as possible as they were presented on the computer screen. For the Inhibit condition (18 trials), children were instructed to name only characters with happy faces and to suppress naming for characters with sad faces. For the final, switching condition (15 trials), children were required to alternate their responses in accordance with a cue; characters wearing hats were to be named by their shapes and characters without hats by their color. Response times and accuracy were coded in Noldus, with inter-rater reliability being 94 and 99%

for each respectively. Dependent variables were the mean reaction time for accurate baseline color naming trials, the proportion of correct inhibit trials and the proportion of correct switch trials.

Trails-Preschool (Espy and Cwik, 2004) was used to assess cognitive flexibility. The task was presented as a story about a family of dogs. During a baseline condition (Trails-P:A), children were asked to stamp the dogs in order of size as quickly as possible. During the subsequent, switching phase of the task (Trails-P:B), children were requested to stamp the dogs and their corresponding bones—also ordered by size—in an alternating sequence. When the child made an incorrect response, he/she was prompted to repeat the response until correct. Performance was coded in Noldus; inter-rater reliability was 99% for response times and 95% for accuracy. Dependent variables included mean reaction time for correct responses during baseline condition A and an efficiency score, computed as the correct responses/total responses for condition B.

The *Visual Matching Test* from the BIA (Woodcock et al., 2001a) was selected as a direct measure of processing speed. In the first segment of the task, children were timed as they pointed to matching shapes as quickly as possible. Following this, they were provided a pencil and asked to circle matching digits as quickly as possible within a 3 min window. Published test-retest reliability is adequate ($r = 0.80$) in the 2–7 years age range.

Mathematics achievement

The Test of Early Mathematics Ability -3 (TEMA-3; Ginsburg and Baroody, 2003) was administered at each follow-up point to assesses children's rudimentary knowledge of numeric concepts, including magnitude comparison, non-verbal addition and subtraction, cardinality, part-whole relationships, mathematic symbol recognition, and counting. The TEMA-3 shows high internal ($\alpha = 0.92 - 0.96$) and test-retest reliability ($r = 0.82 - 0.93$).

The *Applied Problems* subtest from the Woodcock-Johnson Tests of Achievement-III was used to assess children's early mathematical problem-solving abilities at after age 3. The task includes story and picture-based mathematical problems. Test-retest reliability in the younger age ranges is 0.92.

Verbal ability

The *Verbal Comprehension* subtest from the BIA (Woodcock et al., 2001a) was used as a measure of language proficiency. The subtest has four components: picture vocabulary, synonyms, antonyms, and verbal analogies. Test-retest reliability in this age range is high ($r = 0.93$).

ANALYTIC OVERVIEW

Variable distributions were examined for skewness and kurtosis prior to analysis, with outliers trimmed to within 3 *SD* of the mean. Response times also were log transformed, given evidence for significant skew. All models were constructed in MPLUS version 7.11 (Muthen and Muthen, 2012). **Figure 1** describes the model of EC and processing speed, which initially was examined at each individual study age point. As shown, the Visual Matching subtest score was used as a statistical "anchor" for the processing speed factor, as it is a well-used, standardized measure of processing speed. Three other dependent measures from the baseline executive task conditions, namely mean response time for the Shape School baseline naming condition, mean response time for the Big-Little non-conflict trials, and Trails-P:A mean response times, were loaded onto this factor as processing speed measures. These response time variables were reverse-scaled in all presented models to enhance interpretability. In addition, all EC indicators were loaded onto the processing speed factor, thereby allowing any of the variance that EC conditions shared with the less complex, baseline processing conditions to be captured by the processing speed latent. To account for the variance in executive tasks that was shared with language, we used the only language assessment available in this study, the Verbal Comprehension subtest score from the BIA. Given that the Visual Matching and Verbal Comprehension subtests were administered only at study entry and exit, performance at the age point closest to executive task administration was used in

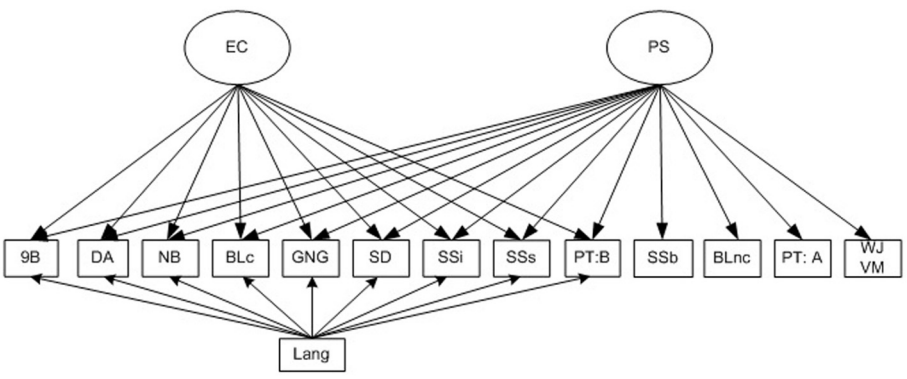


FIGURE 1 | Model of the relations between executive control and processing speed tested at different preschool age points. EC, Executive Control; PS, Processing Speed; 9B, Nine Boxes maximum correct run; DA, Delayed Alternation score; NB, Nebraska Barnyard trials correct; BLc, Big-Little conflict trial accuracy; SD, Snack Delay movement score; SSi, Shape School Inhibit Accuracy; SSs, Shape School Switch accuracy; PT:B, Preschool Trails: B efficiency; SSb, Shape School Baseline naming response time; BLnc, Big-Little non-conflict trail response time; PT:A, Preschool Trails: A response time; WJ VM, Woodcock-Johnson III Visual Matching subtest score; Lang, Woodcock-Johnson III Verbal Comprehension subtest score.

these models. Using an approach similar to Lee et al. (2013), all residuals from the executive measures also were regressed on the Verbal Comprehension task, thus covarying for differences in language proficiency at the manifest level. Finally, executive task conditions were cross-loaded with an EC latent, which captured all of the residual shared variance between manifest executive tasks that was not accounted for by processing speed or the language covariate. All correlations between the latent factors and the latent factors and language proficiency were set to 0, as is common in multidimensional measurement models. Conceptually, this parameterization means that the model describes the contributions of the latent variables to manifest task performance if these variables are assumed to be orthogonal at the construct level. Where dependent measures had been extracted from the same task (e.g., Shape School baseline, Inhibit, and Switch conditions), their residuals were allowed to co-vary on the basis that (1) without accounting for their shared method variance, dependent variables extracted from the same task may have shown spuriously inflated loadings on the latent constructs and thereby clouded understanding of how the latent variables each contribute to task performance, and (2) initial analyses suggested significant improvement in model fit when these residuals were allowed to correlate.

After performing this descriptive analysis of the EC—processing speed overlap in each age group, we extended the analysis to more formally assess statistical changes in the strength of the EC factor loadings over time. This involved combining the models for each age point into a single model and then iteratively constraining the factor loadings to be equal at all age points. Where equality constraints caused a reduction in overall model fit, as evaluated with a chi-squared difference test (Kline, 2011), the loading was freed at one or more age points and the model was re-evaluated.

In the final stage of analysis, we examined the relation of EC and processing speed at each age to mathematics performance. As shown in **Figure 2**, for each independent age point, TEMA-3 and WJ-III Applied Problem subtest scores from the same time point and then every successive assessment point were regressed on processing speed, EC and on the language covariate.

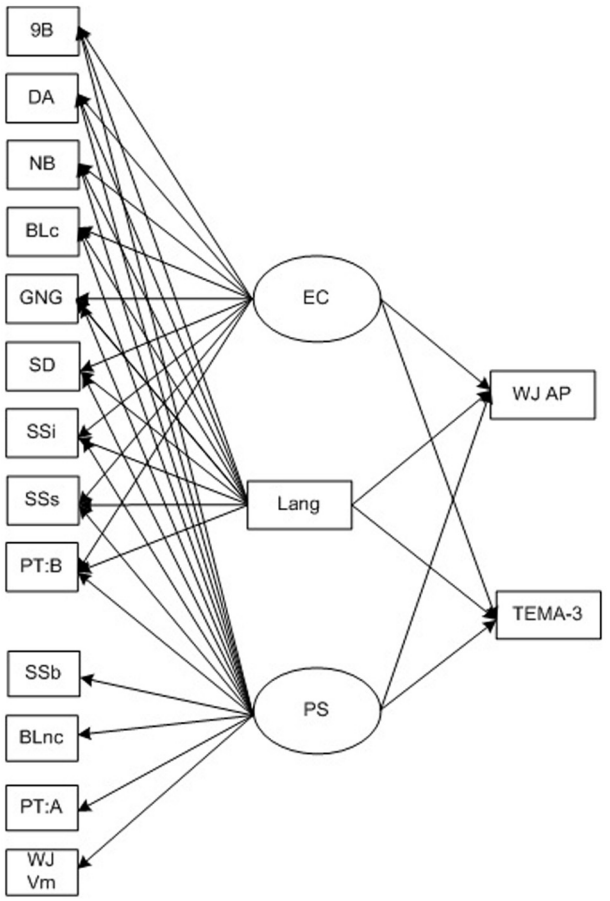


FIGURE 2 | Model of executive control, processing speed and language achievement as predictors of mathematics achievement over the preschool period. EC, Executive Control; PS, Processing Speed; 9B, Nine Boxes maximum correct run; DA, Delayed Alternation score; NB, Nebraska Barnyard trials correct; BLc, Big-Little conflict trial accuracy; SD, Snack Delay movement score; SSi, Shape School Inhibit Accuracy; SSs, Shape School Switch accuracy; PT:B, Preschool Trails: B efficiency; SSb, Shape School Baseline naming response time; BLnc, Big-Little non-conflict trail response time; PT:A, Preschool Trails: A response time; WJ VM, Woodcock-Johnson III Visual Matching subtest; Lang, Woodcock-Johnson III Verbal Comprehension subtest score; WJ AP, Woodcock-Johnson III Applied Problems subtest score.

RESULTS

DESCRIPTIVE PROFILE OF PERFORMANCE ON EXECUTIVE TASKS ACROSS THE PRESCHOOL PERIOD

Table 1 presents descriptive statistics for executive, processing speed and language proficiency measures, as well as the correlations between these tasks at different age points. For Shape School baseline, Big-Little non-conflict and Trails-P:A response times, higher scores reflect slower speed and, therefore, worse performance, whereas scores for all other tasks are positively scaled. Note that in many cases, correlations between executive and processing speed measures were as robust as correlations among the executive tasks themselves, highlighting the interrelations among these putative dimensions of cognition. Children's accuracy on the executive task conditions increased dramatically with age, as did the speed of their responses on non-executive task conditions. A multivariate ANOVA with age group as a predictor supported this pattern of improvement with age, with all univariate effects also significant, $F_{(36,2869.7)} = 38.54$, Wilk's $\lambda_{(36)} = 0.31$, $p < 0.001$. Independent of the multivariate effect of age, there was no significant overall effect of study entry cohort, suggesting little impact of repeated testing on children's executive task performance, $F_{(36,2869.7)} = 1.29$, Wilk's $\lambda_{(36)} = 0.95$, $p = 0.11$.

STRUCTURAL RELATIONS BETWEEN PROCESSING SPEED AND EXECUTIVE CONTROL OVER THE COURSE OF THE PRESCHOOL PERIOD

Table 2 presents a summary of the models for each separate study follow-up point. Specifically, for each separate age, the table shows the standardized coefficients for the model described in **Figure 1**, including the factor loadings for all dependent variables loaded on the processing speed and EC factors, as well as the

Table 1 | Descriptive statistics and correlations between executive control, processing speed, and language measures across the preschool period.

Measure	*M* (*SD*)	1	2	3	4	5	6	7	8	9	10	11	12	13
AGE 3 YEARS (*n* = 228)														
Nine boxes maximum correct run	4.32 (1.60)													
Delayed alternation summary score	−0.37 (3.96)	0.09												
Nebraska barnyard summary score	2.88 (1.57)	0.13*	0.19**											
Big little % correct conflict trials	0.29 (0.29)	0.08	0.23***	0.30***										
Go/No-Go *d*′	0.26 (0.75)	0.03	0.09	0.14*	0.18**									
Snack delay epochs with no movement	12.21 (11.42)	0.07	0.04	0.26***	0.24***	0.23***								
Shape school inhibit % correct	0.38 (0.42)	−0.02	0.18*	0.27***	0.36***	0.28***	0.26***							
Shape school switch % correct	0.26 (0.26)	0.03	0.15*	0.58***	0.25**	0.15	0.17*	0.50***						
Preschool trails B efficiency	0.70 (0.19)	0.02	0.07	0.33***	0.20**	0.15*	0.31***	0.27***	0.28***					
Big little non-conflict response time	3.21 (8.12)	−0.07	−0.05	−0.17*	−0.01	−0.09	−0.16*	−0.12	−0.08	−0.06				
Shape school A response time	2.30 (2.09)	0.00	−0.12	−0.34***	−0.23**	−0.23**	−0.24***	−0.14	−0.19*	−0.16*	0.24***			
Preschool trails A response time	7.51 (6.18)	−0.14*	−0.18**	−0.24***	−0.17*	−0.10	−0.04	−0.14	−0.17*	0.22**	−0.02	0.24***		
Visual matching study entry	413.45 (10.70)	0.07	0.10	0.41***	0.30***	0.16*	0.33***	0.24***	0.28***	0.19*	−0.16*	−0.32***	−0.22**	
Verbal comprehension study entry	436.60 (6.92)	0.05	0.17**	0.46***	0.23***	0.15*	0.26***	0.37***	0.40***	0.22**	−0.11	−0.22**	−0.26***	0.31***
AGE 3.75 YEARS (*n* = 276)														
Nine boxes maximum correct run	4.92 (1.71)													
Delayed alternation summary score	2.61 (4.85)	0.09												
Nebraska barnyard summary score	4.90 (2.07)	0.14*	0.19**											
Big little % correct conflict trials	0.64 (0.34)	0.08	0.22***	0.43***										
Go/No-Go *d*′	1.38 (1.01)	0.11	0.11	0.25***	0.27***									
Snack delay epochs with no movement	17.96 (10.93)	0.08	0.10	0.20***	0.24***	0.30***								
Shape school inhibit % correct	0.75 (0.37)	0.05	0.12*	0.33***	0.36***	0.42***	0.25***							
Shape school switch % correct	0.57 (0.29)	0.14*	0.17**	0.35***	0.51***	0.33***	0.22***	0.48***						
Preschool trails B efficiency	0.82 (0.13)	0.04	0.07	0.18**	0.13*	0.09	0.10	0.28	0.20					
Big little non-conflict response time	1.59 (0.83)	−0.04	−0.03	−0.16**	0.11	−0.21***	−0.11	−0.13*	−0.10	−0.05				
Shape school A response time	1.26 (0.60)	−0.06	−0.13	−0.26***	−0.18**	−0.31***	−0.16**	−0.33***	−0.22***	−0.02	0.30***			
Preschool trails A response time	4.34 (3.52)	−0.10	−0.12	−0.36***	−0.27***	−0.31***	−0.28***	−0.16	−0.21***	0.04	0.08	0.19**		
Visual matching study entry	416.90 (13.28)	0.02	0.18**	0.28***	0.23***	0.23***	0.13*	0.19**	0.37***	−0.08	−0.10	−0.20**	−0.21**	
Verbal comprehension study entry	438.13 (7.74)	0.09	0.06	0.32***	0.30***	0.26***	0.22***	0.31***	0.42***	0.06	−0.11	−0.20***	−0.34***	0.48***

(Continued)

Table 1 | Continued

Measure	*M* (*SD*)	1	2	3	4	5	6	7	8	9	10	11	12	13
AGE 4.5 YEARS (*n* = 319)														
Nine boxes maximum correct run	5.23 (1.79)													
Delayed alternation summary score	5.99 (5.51)	0.09												
Nebraska barnyard summary score	6.98 (2.60)	0.17**	0.23***											
Big little % correct conflict trials	0.85 (0.21)	0.06	0.14*	0.34***										
Go/No-Go *d'*	2.30 (0.82)	0.02	0.18**	0.25***	0.20***									
Snack delay epochs with no movement	21.85 (9.40)	0.16**	0.23***	0.29***	0.16**	0.31***								
Shape school inhibit % correct	0.92 (0.18)	0.00	0.12*	0.21***	0.16**	0.32***	0.23***							
Shape school switch % correct	0.73 (0.24)	0.07	0.18**	0.37***	0.40***	0.33***	0.30***	0.29***						
Preschool trails B efficiency	0.86 (0.12)	0.09	0.16**	0.19***	0.13*	0.14*	0.16**	0.16**	0.22***					
Big little non-conflict response time	1.28 (0.49)	−0.10	−0.20***	−0.20***	0.14*	−0.22***	−0.20***	−0.18**	−0.15*	0.01				
Shape school A response time	1.02 (0.38)	−0.07	−0.13*	−0.13*	−0.08	−0.22***	−0.20***	−0.07	−0.12*	−0.03	0.34***			
Preschool trails A response time	2.21 (1.81)	−0.09	−0.17**	−0.38***	−0.18**	−0.30***	−0.30***	−0.13*	−0.28***	0.10	0.35***	0.38***		
Visual matching study exit	454.77 (11.48)	0.07	0.18**	0.34***	0.30***	0.25***	0.22***	0.18**	0.24***	0.08	−0.15*	−0.26***	−0.38***	
Verbal comprehension study exit	463.23 (12.76)	0.10	0.12*	0.45***	0.27***	0.27***	0.20***	0.16**	0.34***	0.00	−0.27***	−0.18**	−0.37***	0.44***
AGE 5.25 YEARS (*n* = 363)														
Nine boxes maximum correct run	5.73 (1.87)													
Delayed alternation summary score	7.53 (5.81)	0.11*												
Nebraska barnyard summary score	8.91 (2.41)	0.17**	0.14											
Big little % correct conflict trials	0.95 (0.08)	0.15**	0.11	0.38***										
Go/No-Go *d'*	2.74 (0.49)	0.17**	0.24	0.28***	0.25***									
Snack delay epochs with no movement	25.51 (9.85)	0.23***	0.03	0.24*	0.21***	0.20***								
Shape school inhibit % correct	0.96 (0.10)	0.03	0.05	0.13*	0.17***	0.09	0.15**							
Shape school switch % correct	0.84 (0.17)	0.12*	0.07	0.28***	0.28***	0.23***	0.19***	0.23***						
Preschool trails B efficiency	0.90 (0.11)	0.09	0.13*	0.17**	0.19***	0.11*	0.12*	0.12*	0.22***					
Big little non-conflict response time	1.12 (0.33)	−0.10	−0.11*	−0.25***	−0.05	−0.19***	−0.17**	−0.06	−0.17**	−0.06				
Shape school A response time	0.88 (0.25)	−0.02	−0.06	−0.16**	0.04	−0.15*	−0.09	0.05	−0.12*	0.04	0.21***			
Preschool trails A response time	1.28 (0.68)	−0.07	−0.13*	−0.31***	−0.22***	−0.29***	−0.11*	−0.02	−0.03	0.10	0.19***	0.24***		
Visual matching study exit	456.18 (11.99)	0.06	0.05	0.34***	0.26***	0.27***	0.13*	0.05	0.30***	0.05	−0.15**	−0.23***	−0.33***	
Verbal comprehension study exit	464.28 (13.12)	0.03	−0.02	0.53***	0.34***	0.28***	0.14*	0.15**	0.27***	0.13*	−0.18***	−0.14**	−0.24***	0.45***

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. Response times presented in seconds, where a longer response indicates slower speed. The Visual Matching score is a Rasch-based, positively scaled w score (Woodcock et al., 2001a).

Table 2 | Summary of standardized coefficients for multidimensional model by age group.

	λ (S.E) on processing speed factor				β (S.E) on Woodcock–Johnson III language comprehension				λ (S.E) on executive control			
	3	**3.75**	**4.5**	**5.25**	**3**	**3.75**	**4.5**	**5.25**	**3**	**3.75**	**4.5**	**5.25**
Nine boxes maximum correct run	0.11	0.16	0.14	0.16*	0.03	0.06	0.06	0.01	−0.09	0.01	0.07	0.31***
	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.10)	(0.09)	(0.08)	(0.07)
Delayed alternation correct-incorrect runs	0.21**	0.32**	0.28***	0.26***	0.15***	−0.01	0.03	−0.06	0.02	0.08	0.22**	0.20*
	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)	(0.05)	(0.09)	(0.09)	(0.08)	(0.08)
Nebraska barnyard summary score	0.57***	0.52***	0.31***	0.32***	0.40***	0.21***	0.39***	0.48***	−0.58*	0.18**	0.32***	0.31***
	(0.08)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.05)	(0.04)	(0.27)	(0.11)	(0.07)	(0.07)
Big-Little conflict trials % correct	0.48***	0.51***	0.11	0.17**	0.18***	0.21***	0.31***	0.32***	0.12	0.25	0.39***	0.44***
	(0.07)	(0.12)	(0.08)	(0.07)	(0.07)	(0.06)	(0.06)	(0.05)	(0.12)	(0.19)	(0.07)	(0.07)
Go/No-Go *d'*	0.34***	0.40***	0.36***	0.39***	0.11	0.18***	0.31***	0.20***	0.18	0.30**	0.31***	0.26**
	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)	(0.05)	(0.11)	(0.10)	(0.07)	(0.08)
Snack delay movement score	0.40***	0.31***	0.39***	0.19**	0.20***	0.15*	0.17**	0.08	−0.12	0.17	0.31***	0.41***
	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.09)	(0.11)	(0.07)	(0.08)
Shape school condition B % correct	0.43***	0.21***	0.19*	0.03	0.34***	0.26***	0.11	0.16**	0.23	0.85*	0.37***	0.36**
	(0.08)	(0.12)	(0.08)	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)	(0.15)	(0.38)	(0.10)	(0.07)
Shape school condition D % correct	0.42***	0.46***	0.26**	0.19	0.35***	0.35***	0.30***	0.25***	−0.30*	0.35	0.62***	0.37***
	(0.10)	(0.15)	(0.08)	(0.07)	(0.07)	(0.06)	(0.06)	(0.05)	(0.12)	(0.23)	(0.08)	(0.07)
Preschool trails B efficiency	0.32***	0.12**	0.07	0.04	0.29***	0.08	0.04	0.16**	−0.06	0.32	0.38***	0.37***
	(0.08)	(0.16)	(0.09)	(0.09)	(0.07)	(0.07)	(0.06)	(0.05)	(0.10)	(0.18)	(0.08)	(0.07)
Big little non-conflict Trial *M* response time	0.27**	0.31***	0.49***	0.37***								
	(0.08)	(0.08)	(0.06)	(0.07)								
Shape school condition A *M* response time	0.50***	0.47***	0.54***	0.39***								
	(0.07)	(0.08)	(0.06)	(0.07)								
Preschool trails A response time	0.24***	0.43***	0.70***	0.62***								
	(0.08)	(0.08)	(0.06)	(0.07)								
Woodcock-Johnson visual matching score	0.49***	0.25**	0.36***	0.37***								
	(0.07)	(0.07)	(0.07)	(0.07)								

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. All response times are reverse scored for interpretability.*

regression coefficients for executive tasks regressed on the language proficiency covariate. At age 3 years, most executive tasks loaded significantly on processing speed ($\lambda = 0.21-0.46$, $p < 0.05$), the exception being Nine Boxes, $\lambda = 0.11$, $p = 0.16$; Model $\chi^2_{(54)}$ 61.68, $p = 0.22$; CFI = 0.98; RMSEA = 0.03. The majority of executive measures also showed significant relations with the language covariate, although these associations were higher for tasks with more verbal content (i.e., Shape School, Nebraska Barnyard). Very few measures loaded on the independent EC factor ($\lambda = -0.08-0.23$;$p > 0.05$). The exceptions were Nebraska Barnyard ($\lambda = -0.58$, $p = 0.03$) and the Shape School Switch condition ($\lambda = -0.30$, $p = 0.01$), which showed negative loadings. Despite the low loadings on the EC latent, a chi-squared difference test indicated that the model incorporating the EC factor was a significant improvement over a model where the EC loadings were set to 0, $\Delta\chi^2 = 38.6$ (9), $p < 0.001$, although this fit may have been driven by a particularly large increase in the explained variance for Nebraska Barnyard when the EC latent was included ($R^2\Delta = 0.37$).

In the model for the 3.75 age group, the residual variance for Shape School Inhibit was negative, leading to a non-positive definite solution. Once the non-significant residual covariance between Shape School Inhibit accuracy and Shape School Switch accuracy was set to 0, the model converged and was positive definite, although the removal of this residual covariance results in a model that is not directly comparable to models for other age points. All measures loaded significantly on the processing speed latent, $\lambda = 0.16-0.52$, $p < 0.05$; Model $\chi^2_{(55)} = 118.59$, $p < 0.001$; CFI = 0.90; RMSEA = 0.06. Similarly, children's performance on most of the measures, with the exception of Nine Boxes, Delayed Alternation and Trails-P: B, was related to their language proficiency ($\beta = 0.13-0.35$, $p < 0.05$). After accounting for variance shared with processing speed and the language covariate, Nebraska Barnyard, Shape School switch and Go/No-Go loaded significantly and positively on EC. Although model fit statistics indicated that the model provided only an adequate fit to the data, it still provided significantly better fit than a model that did not incorporate an EC latent, $\Delta\chi^2 = 24.80$ (9), $p = 0.003$, with the increase in explained variance being greatest for the Go/No-Go task ($R^2\Delta = 0.27$).

At age 4.5 years, the majority of executive tasks cross-loaded on EC ($\lambda = 0.22-0.62$, $p < 0.01$; Model $\chi^2_{(54)} = 122.8$, $p < 0.001$; CFI = 0.91; RMSEA = 0.06), the exception being Nine Boxes. Most measures also loaded significantly on the processing speed factor. Nebraska Barnyard, Big-Little, Go/No-Go, Snack Delay, and the Shape School Switch condition also showed significant relations with the language covariate. The model including the latent EC factor was a significant improvement over a model where loadings on the EC factor were set to 0, $\Delta\chi^2 = 41.34$ (9), $p < 0.001$. The R^2 values for the individual EC tasks also increased by 2–19% with the addition of the EC latent.

Finally, at age 5.25 years, all EC tasks showed significant loadings of similar magnitude on the EC latent ($\lambda = 0.20-0.38$, $p < 0.05$; Model $\chi^2_{(54)} = 103.37$, $p < 0.001$; CFI = 0.93; RMSEA = 0.05), although Shape School Inhibit, Shape School Switch and Trails-P: B no longer loaded significantly on processing speed. Again, the fit of this model was a significant improvement over a model where the loadings on the EC factor were set to 0, $\Delta\chi^2 = 39.48$ (9), $p < 0.001$ [1], the R^2 values for the manifest variables increasing by 1–20%.

METRIC INVARIANCE OF FACTOR LOADINGS OVER THE COURSE OF THE PRESCHOOL PERIOD

Taken together, the above findings suggest a gradual, age-related increase in the strength and consistency of executive task loadings on the separate EC factor. To more formally evaluate whether these apparent changes in factor loadings were statistically significant, a longitudinal metric invariance analysis was conducted. A combined model, which included the EC and processing speed factors for all four age points, provided a poor fit to the data, even allowing for residual autocorrelations between measures administered at directly successive age points, $\chi^2 = 1550.47$ (1124), $p < 0.001$, CFI = 0.88, RMSEA = 0.03. The majority of factor loadings for EC at age 3 also could not be set equivalent with loadings at the later age points without a significant reduction in model fit. Exceptions were Shape School switch accuracy and Snack Delay, which could be set equivalent between ages 3 and 3.75. By age 3.75, most of the loadings of executive tasks on EC could be constrained equal to those at ages 4.5 and 5.25 years, although Snack Delay and Go/No-Go could not. Finally, the only task loading for EC that could not be constrained to equality between ages 4.5 and 5.25 years was Big- Little, overall $\Delta\chi^2 = 45.88(39)$, $p = 0.09$. Correlations between the EC factors also increased from $\beta = 0.34$ to.45 for EC at age 3 with EC at later ages to $\beta = 0.84$ for EC at age 4.5 with EC at 5.25 years, all p's < 0.001, suggesting increasing stability in the distinct EC factor over time.

RELATIONS OF PROCESSING SPEED, LANGUAGE AND EXECUTIVE CONTROL TO MATHEMATICS ACHIEVEMENT

Table 3 shows the relations of EC, processing speed and the language proficiency covariate to mathematics achievement both at simultaneous and follow-up age points (see **Figure 2** for a description of the model tested independently for each assessment point). Processing speed and language proficiency were robustly correlated with TEMA-3 and WJ-III Applied Problems performance at all time points. However, latent EC at 3 years did not predict mathematics achievement. In contrast, higher latent EC at 3.75 years was associated with higher concurrent TEMA-3 and Applied Problems performance, independent of processing speed and language proficiency, $\chi^2_{(76)} = 149.24$, $p < 0.001$; CFI = 0.93; RMSEA = 0.06. Similarly, EC at 3.75 was

[1] Given that an independent EC factor was evident at age 5.25, we proceeded to test a two-factor EC structure with separate working memory and inhibitory control factors. A model with Nine Boxes, Delayed Alternation, and Nebraska Barnyard loaded on a working memory factor and other tasks loaded on an Inhibitory control factor did not provide better fit to the data, $\Delta\chi^2 = 0.26_{(1)}$, $p = 0.61$ and the correlation between factors was high ($r = 0.91$). Similarly, a model where Shape School Switch and Trails-P: B were loaded on the inhibitory control factor was not superior to the unitary model, $\Delta\chi^2 = 0.003_{(1)}$, $p = 0.96$. In keeping with previous research in this age group, a unitary factor was the preferred model of EC, even when the processing speed and language requirements of executive tasks were accounted for.

Table 3 | Summary of relations of processing speed, language, and executive control to mathematics outcomes across the preschool period.

	β (S.E) on processing speed factor				β (S.E) on Woodcock–Johnson III language comprehension				β (S.E) on executive control			
	3	3.75	4.5	5.25	3	3.75	4.5	5.25	3	3.75	4.5	5.25
TEMA-3 STANDARD SCORE												
3 Years	0.48***				0.44***				−0.12			
	(0.06)				(0.06)				(0.09)			
3.75 Years	0.47***	0.71***			0.51***	0.30***			−0.16	0.18*		
	(0.07)	(0.05)			(0.05)	(0.05)			(0.10)	(0.08)		
4.5 Years	0.45***	0.59***	0.60***		0.40***	0.32***	0.43***		−0.19	0.23*	0.14	
	(0.07)	(0.07)	(0.07)		(0.06)	(0.06)	(0.05)		(0.10)	(0.09)	(0.11)	
5.25 Years	0.35***	0.61***	0.46***	0.47***	0.38***	0.24***	0.46***	0.48***	−0.11	0.15	0.19*	0.37***
	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.10)	(0.10)	(0.08)	(0.06)
WOODCOCK-JOHNSON APPLIED PROBLEMS SCORE												
3.75 Years	0.34***	0.50***			0.60***	0.45***			−0.06	0.31***		
	(0.07)	(0.07)			(0.05)	(0.06)			(0.09)	(0.08)		
4.5 Years	0.43***	0.52***	0.51***		0.45***	0.38***	0.46***		−0.03	0.40***	0.35***	
	(0.06)	(0.08)	(0.08)		(0.06)	(0.06)	(0.05)		(0.10)	(0.08)	(0.10)	
5.25 Years	0.35***	0.59***	0.47***	0.48***	0.39***	0.31***	0.47***	0.50***	−00.09	0.21*	0.30***	0.27***
	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.04)	(0.11)	(0.10)	(0.07)	(0.07)

*$p < 0.05$; ***$p < 0.001$. Separate models were constructed for each EC assessment and mathematics outcome age.

associated with higher TEMA-3 and Applied Problems performance at age 4.5 years, $\chi^2_{(76)} = 151.07$, $p < 0.001$; CFI = 0.93; RMSEA = 0.06. Higher latent EC at 4.5 years also was associated with higher Applied Problems performance both concurrently (Model $\chi^2_{(76)} = 184.70$, $p < 0.001$; CFI = 0.91; RMSEA = 0.07) and at the 5.25 year follow-up, Model $\chi^2_{(76)} = 181.36$, $p < 0.001$; CFI = 0.92; RMSEA = 0.07. Finally, EC at age 5.25 was independently related to TEMA-3 and Applied Problems performance at the same 5.25 year age point, Model $\chi^2_{(76)} = 153.88$, $p < 0.001$; CFI = 0.94; RMSEA = 0.05. Findings for all models were similar when the effect of study entry cohort was considered.

DISCUSSION

The marked overlap in adult performance on measures of EC and general processing speed has triggered debate regarding the validity of EC as a distinct, independent dimension of cognition (Rabitt, 1997; Salthouse, 2005). In early childhood, a period when both EC and processing speed improve dramatically, such issues related to the construct validity of EC intersect with questions regarding the nature of EC development and the potentially cascading impact of advancements in basic processing fluency on higher-order cognition. The aims of this study were to examine the overlap between measures of EC and processing speed at different preschool age points and test the predictive utility of EC in relation to children's mathematics achievement after accounting for the processing demands of early executive tasks. Findings indicate that EC and processing speed are highly intertwined in early childhood to the extent that their impact on executive task performance at age 3 years could not be cleanly parsed. As children age through the preschool period, EC progressively differentiates from processing speed, becomes more stable, and shows independent predictive relations with mathematics achievement. Not only does this study shed some light on the psychometric characteristics of early EC tasks, but it also provides insight into the developmental mechanisms that might facilitate executive proficiency in early childhood.

Cascade models posit that increases in processing speed facilitate the development of higher-order executive skills (e.g., Fry and Hale, 1996). Support for this hypothesis derives from studies showing that processing speed mediates the relation of age to EC. Yet mediation analyses cannot reveal potential changes in the interplay of processing speed and EC over development. Consistent with the cascade model, this study does suggest that processing speed contributes substantially to children's performance on executive tasks. Nonetheless, the study also provides evidence that there are qualitative shifts in the interface of these processes over time. All of the EC factor loadings at age 3 were negative or non-significant, indicating that a general processing speed factor is able to explain all of the overlap in children's performance and that any residual variance after the processing and language demands of executive tasks are accounted for is largely specific to individual tasks. It is possible that at this young age, children draw to a greater extent on baseline processing and language skills to perform executive tasks, meaning that variability in executive performance is driven primarily by individual differences in these skills. A second possibility is that measures are not sufficiently sensitive to distinct dimensions of cognition because of the high level of within-person variability in young children's motivation and fatigue, implying that the processing speed factor reflects a broader, non-specific characteristic, such as task engagement or attention. A final possibility is that processing speed and EC are too tightly intertwined and co-dependent in this young age group. Even basic processing of shapes and colors may to some extent involve effortful cognitive control because children have not yet mastered these concepts, making it difficult

to disentangle the unique roles of EC and processing speed in behavioral performance.

The relations between executive tasks increased over time, with some tasks beginning to load positively on a separate EC factor by age 3.75, although tasks also continued to load consistently on the processing speed factor through the preschool period. Quicker information processing may provide a platform for EC by freeing up higher-order resources, enabling children to hold more rules or situational requirements in working memory. Processing speed may also facilitate inhibition of motor or vocal responses because activation of inhibitory control networks can occur more quickly. This tight coupling between general processing speed and EC may help to explain why deficits in executive task performance characterize so many psychological disorders and why childhood traumatic brain injury to any area of the brain is associated with lower EC task performance (Jacobs et al., 2011). Disruption to cortical circuitry, regardless of its area in the brain, is likely to slow neural processing and transmission, with consequent bottom-up effects on EC. Even in older children, processing speed appears to mediate a substantial part, although not all, of the relation between age and complex working memory task performance (Bayliss et al., 2005; Fry and Hale, 1996). Recent studies also suggest that slow processing speed explains much of the deficit in working memory and inhibitory control performance in children with ADHD relative to their typically-developing peers (Lijffijt et al., 2005; Karulunas and Huang-Pollack, 2013).

At all age points, language proficiency also predicted residual variance in executive performance that was not explained by processing speed. The strong links between language abilities and EC often are framed in terms of social interactions and cultural tools, which theoretically create a symbol system that children can use to represent concepts or rules or to engage in internalized speech that allows them to self-regulate (Vygotsky, 1978). Perhaps its unique relation to language through these symbolic codes serves in part to differentiate EC from more general processing speed.

By the end of the preschool period, the common requirements of task conditions that had been manipulated to capture EC clearly diverged from processing speed and formed a coherent latent construct that was relatively stable from age 4.5 to 5.25 years. From a psychometric perspective, these findings provide evidence for the divergent validity and sensitivity of executive measures from about age 4 years. The extraction of shared task variance above and beyond that associated with processing speed and language allows for greater confidence that cognition incorporates a distinct, top-down control system, which is engaged specifically when tasks include demands for cognitive flexibility, the on-line maintenance and updating of task-relevant information, or the inhibition of a prepotent response. It should also be noted that some measures appear to be stronger indicators of EC than others. Despite their strong basis in animal studies of prefrontal function, Nine Boxes and Delayed Alternation showed lower and somewhat inconsistent correlations with other measures. The combination of processing speed, EC, and language comprehension explained only a small proportion of the variance for these tasks (3–12%). In contrast, Nebraska Barnyard, Big-Little Stroop, Go/No-Go, and Snack Delay showed relatively consistent correlations with each other across the preschool age range, suggesting that they may be more reliable indicators of EC. Collectively, EC, processing speed and the language covariate explained 15–82% of the variance in children's performance on these measures at different ages, whereas the maximum amount of variance explained in studies where our group has modeled EC without accounting for overlap with processing speed at the manifest level is only 57% (Nelson et al., 2014; see Willoughby et al., 2013 for similar findings).

From a theoretical perspective, the age-related divergence of EC and processing speed supports the differentiation hypothesis, where cognitive systems are thought to become progressively specialized over time (Hülür et al., 2011). Functional MRI studies show that, as children's performance on EC tasks improves, neural activation patterns become more focal and localized to regions of the brain that are typically activated when adults perform EC tasks (Durston et al., 2006; Rubia et al., 2006). Bell and Wolfe (2007) found that, in infancy, EEG activity during a working memory task was diffuse across the scalp. In the same group of children at age 4.5 years, however, EEG activity during working memory tasks was localized coherently at frontal electrode sites. There also is increasing development of long-range neural connections across childhood that presumably allow disparate neural systems to communicate more effectively (Fair et al., 2007). The gradual fractionation of EC abilities from processing speed evident in the current study is in line with this movement from a more diffuse activation of neural networks to the functional specialization of cortical circuits that coordinate cognitive control. However, it is also important to note that although EC appeared gradually to differentiate from processing speed, separate inhibition and working memory components of EC were not evident even by the final time point of this preschool study.

Processing speed and language proficiency were strong predictors of children's mathematics performance across the preschool years, whereas latent EC at age 3 years was not related to mathematics achievement once the processing speed and language demands of the EC tasks had been accounted for. Note that we are not suggesting that executive task performance at age 3 years is not a useful predictor of later mathematics achievement. As described in our earlier work, children's performance on many of the EC tasks at age 3 years correlates moderately with their mathematics achievement through the preschool period (Clark et al., 2013). What is clear is that the distinctions between EC and processing speed are not as clear-cut at age 3 and children's general processing speed may in fact drive the correlation between executive task performance and later mathematics. From age 3.75 years, EC did show independent correlations with mathematics achievement over and above individual differences in basic processing abilities and language. These findings provide more support for the construct validity and utility of EC, at least as assessed later in the preschool period. They also provide compelling evidence for the importance of both general processing speed and EC in children's early mathematics acquisition. Processing speed may reflect a central limiting mechanism that constrains or enhances children's ability to quickly retrieve or activate representations such as shapes, words or digits that are essential for mathematics. However, EC likely plays an added

role in allowing for the maintenance and manipulation of these representations, which is essential for on-line mathematics problem solving. It will be important to extend these models to older age groups. Conceivably, the role of EC in mathematics could continue to increase over time. However, it is also possible that increasingly automatic and fluent numeric processing might eventually dampen the requirements for EC as children learn, resulting in differential relations to components of mathematics that have been mastered and those that are not as fluent over time.

It is important to note some limitations of the study. First, it is difficult to obtain pure measures of processing speed and measures of general reaction time may reflect other aspects of performance, including speed-accuracy trade-offs or lapses in attention (Schmiedek et al., 2007). The use of a factor score capturing variance from very different types of tasks was helpful in addressing this issue. Second, while it would have been ideal to construct a factor for language proficiency, constraints on the number of assessments that young children can feasibly complete limited our ability to acquire multiple indicators of language. Finally, in a recent study of EC in school-aged children, reaction times for baseline and executive task conditions could not be separated into distinct factors, whereas accuracy measures did form distinct EC components, highlighting an important influence of the type of indicator chosen on the measurement model for EC (van der Ven et al., 2013). As in most studies of preschoolers, we used accuracy or efficiency measures for the executive conditions of the EC tasks. While unlikely, given the use of varied scoring methods across tasks, is possible that the distinction between EC and processing speed in the later age groups is an artifact of the fact that most of the processing speed indicators were reaction times and most EC indicators were accuracy/efficiency measures.

Despite these limitations, this study clearly adds to the understanding of the nature and importance of EC by demonstrating dynamic changes in the overlap between processing speed and EC in early childhood and a qualitative re-organization of these interfacing processes over time. Early in the preschool period, executive tasks may not be sensitive indicators of an independent EC construct because EC is so intertwined with children's fluency of information processing. As children mature and their processing speed improves, a distinct EC construct plays a greater role in their EC task performance and this EC factor relates independently to children's developing mathematics proficiency. A key message from the study is that there is cause for optimism regarding the potential of specific EC assessment and intervention to address some of the pervasive discrepancies in children's academic readiness. This enthusiasm should be tempered, however, with the recognition of a corpus of psychological research demonstrating that the basic fluency with which children process information is an underpinning platform for intellectual development.

ACKNOWLEDGMENTS

This research was made possible by grants MH065668 and DA023653 from the National Institutes of Health. We acknowledge all the research assistants in the Developmental Cognitive Neuroscience Laboratory-Lincoln who assisted with data collection and coding. Most importantly, we wish to thank all of the children and families who gave up their time to participate in this study. The authors have no financial or commercial relationships that would constitute conflicts of interest.

REFERENCES

Anderson, P. J. (2008). "Towards a developmental model of executive function," in *Executive Functions and the Frontal Lobes*, eds V. Anderson, R. Jacobs, and P. J. Anderson (New York, NY: Psychology Press), 3–21.

Anderson, P. J., and Reidy, N. (2012). Assessing executive function in preschoolers. *Neuropsychol. Rev.* 22, 345–360. doi: 10.1007/s11065-012-9220-3

Bayliss, D. M., Jarrold, C., Baddeley, A. D., Gunn, D. M., and Leigh, E. (2005). Mapping the developmental constraints on working memory span performance. *Dev. Psychol.* 41, 579–597. doi: 10.1037/0012-1649.41.4.579

Bell, M. A., and Wolfe, C. D. (2007). Changes in brain functioning from infancy to early childhood: evidence from EEG power and coherence during working memory tasks. *Dev. Neuropsychol.* 31, 21–38. doi: 10.1207/s15326942dn3101_2

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., et al. (2008). Promoting academic and social-emotional school readiness: the head start REDI program. *Child Dev.* 79, 1802–1817. doi: 10.1111/j.1467-8624.2008.01227.x

Bull, R., Espy, K. A., and Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev. Neuropsychol.* 33, 205–228. doi: 10.1080/87565640801982312

Bull, R., and Johnston, R. S. (1997). Children's arithmetical difficulties: contributions from processing speed, item identification, and short-term memory. *J. Exp. Child Psychol.* 65, 1–24. doi: 10.1006/jecp.1996.2358

Case, R., Kurland, D. M., and Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *J. Exp. Child Psychol.* 33, 386–404. doi: 10.1016/0022-0965(82)90054-6

Chan, R. C. K., Shum, D., Toulopoulou, T., and Chen, E. Y. H. (2008). Assessment of executive functions: review of the instruments and identification of critical issues. *Arch. Clin. Neuropsychol.* 23, 201–216. doi: 10.1016/j.acn.2007.08.010

Clark, C. A. C., Pritchard, V. E., and Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Dev. Psychol.* 46, 1176–1191. doi: 10.1037/a0019672

Clark, C. A. C., Sheffield, T. D., Wiebe, S. A., and Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Dev.* 84, 662–677. doi: 10.1111/j.1467-8624.2012.01854.x

Diamond, A., Barnett, W. S., Thomas, J., and Munro, S. (2007). Preschool program improves executive function. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Diamond, A., Prevor, M. B., Callender, G., and Druin, D. P. (1997). Prefrontal cortex cognitive differences in children treated continuously for PKU. *Monogr. Soc. Res. Child Dev.* 62, 1–206. doi: 10.2307/1166208

Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J. A., Fossella, J. A., et al. (2006). A shift from diffuse to focal cortical activity with development. *Dev. Sci.* 9, 1–20. doi: 10.1111/j.1467-7687.2005.00454.x

Espy, K. A. (1997). The Shape School: assessing executive function in preschool children. *Dev. Neuropsychol.* 13, 495–499. doi: 10.1080/87565649709540690

Espy, K. A. (1999). Executive functioning in preschool children: performance on A-not-B and other delayed response format tasks. *Brain Cogn.* 41, 178–199. doi: 10.1006/brcg.1999.1117

Espy, K. A., and Cwik, M. F. (2004). The development of a trail making test in young children: the TRAILS-P. *Clin. Neuropsychol.* 18, 411–422. doi: 10.1080/138540409052416

Fair, D. A., Dosenbach, N. U., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., et al. (2007). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13507– 13512. doi: 10.1073/pnas.0705843104

Fournier-Vicente, S., Larigauderie, P., and Gaonac'h, D. (2008). More dissociations and interactions within central executive functioning: a comprehensive latent-variable analysis. *Acta Psychol.* 129, 32–48. doi: 10.1016/j.actpsy.2008.04.004

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost

entirely genetic in origin. *J. Exp. Psychol. Gen.* 137, 201–225. doi: 10.1037/0096-3445.137.2.201

Fry, A., and Hale, S. (1996). Processing speed, working memory and fluid intelligence: evidence for a developmental cascade. *Psychol. Sci.* 7, 1467–9280. doi: 10.1111/j.1467-9280.1996.tb00366.x

Fry, A., and Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biol. Psychol.* 54, 1–34. doi: 10.1016/S0301-0511(00)00051-X

Fuhs, M. W., and Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at head start. *Dev. Psychol.* 47, 404–416. doi: 10.1037/a0021065

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. *Dev. Psychol.* 47, 1539–1552. doi: 10.1037/a0025510

Geary, D. C., Hoard, M. K., Nugent, L., and Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: a five-year prospective study. *J. Educ. Psychol.* 104, 206–223. doi: 10.1037/a0025398

Ginsburg, H. P., and Baroody, A. J. (2003). *Test of Early Mathematics Achievement – 3rd Edn*. Austin, TX: Pro-Ed.

Goldman, P. S., Rosvold, H. E., Vest, B., and Galkin, T. W. (1971). Analysis of the delayed alternation deficit produced by dorsolateral prefrontal lesions in the rhesus monkey. *J. Comp. Physiol. Psychol.* 77, 212–220. doi: 10.1037/h0031649

Grudnik, J. L., and Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence* 29, 523–535. doi: 10.1016/S0160-2896(01)00078-2

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Dev.* 61, 653–663. doi: 10.2307/1130951

Hughes, C., Dunn, J., and White, A. (1998). Trick or treat?: uneven understanding of mind and emotion and executive dysfunction in 'hard to manage' preschoolers. *J. Child Psychol. Psychiatry* 39, 981–994. doi: 10.1111/1469-7610.00401

Hughes, C., Ensor, R., Wilson, A., and Graham, A. (2010). Tracking executive function across the transition to school: a latent variable approach. *Dev. Neuropscyhol.* 35, 20–36. doi: 10.1080/87565640903325691

Huizinga, M., Dolan, C. V., and van der Molan, M. W. (2006). Age-related changes in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Hülür, G., Wilhelm, O., and Robitzsch, A. (2011). Intelligence differentiation in early childhood. *J. Individ. Dif.* 32, 170–179.

Jacobs, R., Harvey, A. S., and Anderson, V. (2011). Are executive skills primarily mediated by the prefrontal cortex in childhood? Examination of focal brain lesions in childhood. *Cortex* 47, 808–824. doi: 10.1016/j.cortex.2010.06.002

Kail, R. (1991a). Developmental change in speed of processing during childhood and adolescence. *Psychol. Bull.* 109, 490–501. doi: 10.1037/0033-2909.109.3.490

Kail, R. (1991b). Processing time declines exponentially during childhood and adolescence. *Dev. Psychol.* 27, 259–266. doi: 10.1037/0012-1649.27.2.259

Kail, R. (2002). Developmental change in proactive interference. *Child Dev.* 73, 1703–1714. doi: 10.1111/1467-8624.00500

Kail, R., and Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychol.* 86, 199–225. doi: 10.1016/0001-6918(94)90003-5

Karulunas, S. L., and Huang-Pollack, C. L. (2013). Integrating impairments in reaction time and executive function using a diffusion model framework. *J. Abnorm. Child Psychol.* 41, 837–850. doi: 10.1007/s10802-013-9715-2

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling, 3rd Edn.* New York, NY: Guildford Press.

Kochanska, G., Murray, K. T., and Harlan, E. T. (2000). Effortful control in early childhood: continuity and change, antecedents, and implications for social development. *Dev. Psychol.* 36, 220–232. doi: 10.1037/0012-1649.36.2.220

Kochanska, G., Murray, K. T., Jacques, T. Y., Koenig, A. L., and Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Dev.* 67, 490–507. doi: 10.2307/1131828

Korkman, M., Kirk, U., and Kemp, S. (1998). *NEPSY: A Developmental Neuropsychological Assessment*. Boston, MA: The Psychological Association.

Lee, K., Bull, R., and Ho, R. M. H. (2013). Developmental changes in executive functioning. *Dev. Psychol.* 84, 1933–1953. doi: 10.1111/cdev

LeFevre, J., Berrigan, L., Vendetti, C., Kamawar, D., Bisanz, J., Skwarchuk, S., et al. (2013). The role of executive attention in the acquisition of mathematical skills for children in grades 2 through 4. *J. Exp. Child Psychol.* 114, 243–261. doi: 10.1016/j.jecp.2012.10.005

Lijffijt, M., Kenemans, J. L., Verbaten, M. N., and van Engeland, H. (2005). A meta-analytic review of stopping performance in ADHD: deficient inhibitory motor control? *J. Abnorm. Psychol.* 114, 216–222. doi: 10.1037/0021-843X.114.2.216

McAuley, T., and White, D. A. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *J. Exp. Child Psychol.* 108, 453–468. doi: 10.1016/j.jecp.2010.08.009

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., and Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Muthen, B. O., and Muthen, L. M. (2012). Mplus (Version 7.0). [Computer software]. Los Angeles, CA: Muthen & Muthen.

Nelson, J. M., Sheffield, T. D., Chevalier, N., Clark, C. A., and Espy, K. A. (2014). "Structure, measurement and development of preschool executive control," in *Executive Function in Preschool Children: Integrating Measurement, Neurodevelopment and Translational Research*, eds A. Griffin, L. S. Freund, and P. McArdle (Washington, DC: APA Press), (in press).

Rabitt, P. (1997). "Introduction: methodologies and models in the study of executive function," in *Methodology of Frontal and Executive Function*, ed P. Rabitt (London: Psychology Press), 1–34.

Rose, S. A., Feldman, J. F., and Jankowski, J. J. (2011). Modeling a cascade of effects: the role of speed and executive functioning in preterm/full-term differences in academic achievement. *Dev. Sci.* 14, 1161–1175. doi: 10.1111/j.1467-7687.2011.01068.x

Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E., et al. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Hum. Brain Mapp.* 27, 973–993. doi: 10.1002/hbm.20237

Salthouse, T. A. (1996). The processing speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295X.103.3.403

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology* 19, 532–545. doi: 10.1037/0894-4105.19.4.532

Salthouse, T. A., Atkinson, T. M., and Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *J. Exp. Psychol. Gen.* 132, 566–594. doi: 10.1037/0096-3445.132.4.566

Schmiedek, F., Oberauer, K., Wilhelm, O., Heinz-Martin, S., and Wittman, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *J. Exp. Psychol. Gen.* 136, 414–429. doi: 10.1037/0096-3445.136.3.414

Simpson, A., and Riggs, K. J. (2006). Conditions under which children experience inhibitory difficulty with a "button press" go/no go task. *J. Exp. Child Psychol.* 94, 18–26. doi: 10.1016/j.jecp.2005.10.003

Span, M. M., Ridderinkhof, K. R., and van der Molen, M. W. (2004). Age-related changes in the efficiency of cognitive processing across the life span. *Acta Psychol.* 117, 155–183. doi: 10.1016/j.actpsy.2004.05.005

van der Sluis, S., de Jong, P., and van der Leij, A. (2007). Executive functioning in children and its relations with reasoning, reading and arithmetic. *Intelligence* 35, 427–449. doi: 10.1016/j.intell.2006.09.001

van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., and Leseman, P. P. M. (2013). The structure of executive functions in children: a closer examination of inhibition, shifting and updating. *Br. J. Dev. Psychol.* 31, 70–87. doi: 10.1111/j.2044-835X.2012.02079.x

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge: Harvard University Press.

Wiebe, S. A., Espy, K., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: 1. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575

Wiebe, S. A., Sheffield, T. D., and Espy, K. A. (2012). Separating the fish from the sharks: a longitudinal study of preschool response inhibition. *Dev. Psychol.* 83, 1245–1261. doi: 10.1111/j.1467-8624.2012.01765.x

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., and Espy, K. A. (2011). The structure of executive function in 3-year-olds. *J. Exp. Child Psychol.* 108, 436–452. doi: 10.1016/j.jecp.2010.08.008

Willoughby, M. T., Blair, C., Wirth, R. J., and Greenberg, M. T. (2010). The measurement of executive function at age 3 years: psychometric properties and criterion validity of a new battery of tasks. *Psychol. Assess.* 22, 306–317. doi: 10.1037/a0018708

Willoughby, M. T., Pek, J., and Blair, C. (2013). Measuring executive function in early childhood: a focus on maximal reliability and the derivation of short forms. *Psychol. Assess.* 25, 664–670. doi: 10.1037/a0031747

Wolfe, C. D., and Bell, M. A. (2007). Sources of variability in working memory in early childhood: a consideration of age, temperament, language, and brain electrical activity. *Cogn. Dev.* 22, 431–455. doi: 10.1016/j.cogdev.2007.08.007

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001a). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001b). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 November 2013; accepted: 27 January 2014; published online: 17 February 2014.

Citation: Clark CAC, Nelson JM, Garza J, Sheffield TD, Wiebe SA and Espy KA (2014) Gaining control: changing relations between executive control and processing speed and their relevance for mathematics achievement over course of the preschool period. Front. Psychol. ***5****:107. doi: 10.3389/fpsyg.2014.00107*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Clark, Nelson, Garza, Sheffield, Wiebe and Espy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 16 May 2014
doi: 10.3389/fpsyg.2014.00443

Longitudinal and concurrent links between memory span, anxiety symptoms, and subsequent executive functioning in young children

Laura Visu-Petra[1]*, Oana Stanciu[2], Oana Benga[1], Mircea Miclea[3] and Lavinia Cheie[1]

[1] *Developmental Psychology Lab, Department of Psychology, Babeş-Bolyai University, Cluj-Napoca, Romania*
[2] *Department of Applied Mathematics and Computer Science, University of Ghent, Ghent, Belgium*
[3] *Department of Psychology, Applied Cognitive Psychology Center, Babeş-Bolyai University, Cluj-Napoca, Romania*

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Sebastian B. Gaigg, City University London, UK
Hannah R. Snyder, University of Denver, USA

****Correspondence:***
Laura Visu-Petra, Department of Psychology, Babeş-Bolyai University, Republicii Str. No 37, Cluj-Napoca 400015, Romania
e-mail: laurapetra@psychology.ro

It has been conjectured that basic individual differences in attentional control influence higher-level executive functioning and subsequent academic performance in children. The current study sets out to complement the limited body of research on early precursors of executive functions (EFs). It provides both a cross-sectional, as well as a longitudinal exploration of the relationship between EF and more basic attentional control mechanisms, assessed via children's performance on memory storage tasks, and influenced by individual differences in anxiety. Multiple measures of verbal and visuospatial short-term memory (STM) were administered to children between 3 and 6 years old, alongside a non-verbal measure of intelligence, and a parental report of anxiety symptoms. After 9 months, children were re-tested on the same STM measures, at which time we also administered multiple measures of executive functioning: verbal and visuospatial working memory (WM), inhibition, and shifting. A cross-sectional view of STM development indicated that between 3 and 6 years the trajectory of visuospatial STM and EF underwent a gradual linear improvement. However, between 5 and 6 years progress in verbal STM performance stagnated. Hierarchical regression models revealed that trait anxiety was negatively associated with WM and shifting, while non-verbal intelligence was positively related to WM span. When age, gender, non-verbal intelligence, and anxiety were controlled for, STM (measured at the first assessment) was a very good predictor of overall executive performance. The models were most successful in predicting WM, followed by shifting, yet poorly predicted inhibition measures. Further longitudinal research is needed to directly address the contribution of attentional control mechanisms to emerging executive functioning and to the development of problematic behavior during early development.

Keywords: executive functions, working memory, short term memory span, anxiety, inhibition, shifting, young children

INTRODUCTION

During the past decades, the importance of investigating the early development of executive functions (EFs) has been reinforced by a growing body of evidence linking preschool EFs measures to emerging academic success (see Willoughby et al., 2012a, for a recent review), to social competence during early school years (Ciairano et al., 2007; Razza and Blair, 2009), and also to internalizing and externalizing symptoms (Thorell and Wåhlstedt, 2006; Brocki et al., 2010; Hughes and Ensor, 2011). This endeavor was previously constrained by the limited methodological repertoire allowing researchers to track EF progress across successive developmental periods. Recently, the gap has been addressed by developing a wide range of child-friendly tasks for measuring EF during early development (see Carlson, 2005; Garon et al., 2008 for reviews), with evidence of relatively reliable psychometric properties for this age span (Miller et al., 2012; Willoughby et al., 2012a).

However, the early developmental course and changing structure of executive functioning is not yet fully captured by the limited body of prospective longitudinal data (but see Hughes et al., 2010; Röthlisberger et al., 2012; and Willoughby et al., 2012b, for notable exceptions), most of the research in the field being still cross-sectional. Also, the fundamental prerequisites from the first years of life have not yet been convincingly linked to the intricate nature of later EF, which has been regarded as the most complex form of high-level human cognition (Salthouse, 2005). Moreover, executive control is also determined by, and influential for, emotion-cognition interactions (Pessoa, 2008), which generate stable predispositions in information processing mechanisms (e.g., Pine, 2007), regarded as early cognitive vulnerability markers for a variety of psychopathologies such as internalizing disorders (Ingram and Price, 2010). Further longitudinal studies complementing the limited existing literature (e.g., Riggs et al., 2004; Hughes and Ensor, 2011; Tillman et al., 2013) are necessary

in order to construct true developmental models of how early EF and socio-emotional processes interact to generate problematic behavior and cognitive vulnerabilities to psychopathology.

EARLY EF DEVELOPMENT AND ITS PRECURSORS

With regards to the early developmental trajectory of executive control, initial models argued for the predominant role of one EF, such as *inhibitory control* (Diamond and Gilbert, 1989; Dempster, 1992; Barkley, 1997; Carlson et al., 1998) or *working memory* (WM; Pascual-Leone, 1970; Case, 1985; Morton and Munakata, 2002). A step forward consisted in considering both inhibition and WM as central to EF development (Diamond, 1991; Roberts and Pennington, 1996). The seminal model proposed by Miyake and collaborators (2000) identified three "independent, yet interdependent" EF dimensions: updating of WM representations, inhibition, and shifting. This model was later refined and the identity of inhibition as a distinct factor was questioned. Inhibition subsequently came to be related to common variance in EF tasks (e.g., Friedman et al., 2008; Miyake and Friedman, 2012). The third dimension, *shifting* was defined as the ability to flexibly shift among distinct but related aspects of a given stimulus or task set (Zelazo and Müller, 2002). The tripartite model of EF has been partially confirmed by latent variables analyses conducted in older children samples (Lehto et al., 2003; Huizinga et al., 2006; but see Lee et al., 2012; Van der Ven et al., 2012, for failures to replicate this structure). However, similar studies with preschool children have pointed toward a more unitary structure of EF (Wiebe et al., 2008, 2011; Hughes et al., 2010; Willoughby et al., 2010), although a two-dimensional structure, integrating WM and inhibition as separate yet related factors, was also found (Lerner, 2012; Miller et al., 2012).

Our study was designed to investigate the early developmental interrelations between individual differences in attentional control, memory storage, anxiety symptoms, and subsequent executive functioning (WM, inhibition, shifting) during preschool years. Therefore, we will now review the available evidence on the precursors and subcomponents of these three EF dimensions. The few existing longitudinal studies have generally overlooked how preschool EFs are linked to more basic precursors, such as attentional or memory processes. However, in the literature, there have been some theoretical conjectures regarding these elementary forms of EF. One of the most well-articulated frameworks has been proposed by Garon and collaborators (2008). The authors argue that EF components are built upon simpler cognitive skills and represent the coordination of these basic skills, essentially occurring after the age of three. As a potential candidate, they suggested that the "maturation of attentional capacity forms a foundation for the development of EF abilities during the preschool period, and, in fact, may be the source of common variance underlying various EF skills" (p. 35). Simple span tasks have been proven to rely on individuals' ability to consistently focus and control their attention in order to maintain or suppress information (Engle, 2002) and therefore, might represent an ideal context in which to assess early attentional precursors of EF.

WM processes relate to the updating and active use of temporarily available information. Complementary to this definition, short-term memory (STM) represents the temporarily increased availability of information in memory that may be used to carry out various types of mental tasks (Cowan et al., 1999). The model proposed by Baddeley and Hitch (1974; see also Baddeley, 2000) represents the preferred theoretical framework in which WM development is studied. Various *simple memory span* tasks have been used to measure the two STM storage systems: the phonological loop, and the visuospatial sketchpad. These "slave" systems feed their input into the central executive, a system involved in supervising and adjusting the control of memory contents. Almost all STM measures present a steady *increase* from the preschool years until adolescence (Gathercole et al., 2004; Alloway et al., 2006). *Complex memory span* tasks involve both maintenance and manipulation of information, and are considered measures of WM capacity. The memory components corresponding to the central executive, the phonological loop, and the visuospatial sketchpad appear to resemble the adult tripartite distinction, and to be evident in children as young as four (Alloway et al., 2006). However, it is important to mention that there are strong competing models (Engle et al., 1999; Cowan, 2001; Barrouillet et al., 2004), most of them focusing on the importance of attentional control mechanisms involved in both information storage and processing. Further research on longitudinal interrelations between early aspects of attentional control, memory storage and processing could benefit the integration of the multiple theoretical accounts of WM development.

Two different perspectives could be proposed regarding the involvement of STM processes in WM development and in relation to EF tasks, in general. One of them considers that the active manipulation of information is essential to WM/EF processes (Miyake et al., 2000). Hence, tasks requiring only memory span and which lack this dimension would share only non-executive variance with WM/EF tasks (Lerner, 2012). Another perspective suggests that both simple span and WM tasks share common attentional control demands, and thus their covariance would rely on both executive and non-executive processes. More specifically, WM processes reflect the functioning of the central attention system and its role in the coordination of the systems involved in storage (Garon et al., 2008). The authors argue for the need to conduct longitudinal studies using both complex and simple span tasks in order to "draw conclusions about whether complex WM tasks build upon simpler memory abilities and skills" (p. 40). Beyond its importance for WM development, STM performance could also be predictive for performance in other EF tasks, such as *inhibitory control*. When analyzing the early development of inhibitory control, the focus is mainly on executive inhibitory processes, defined as processes for intentional control or response suppression in the service of higher order or longer-term goals (Nigg, 2000). Friedman and Miyake (2004) empirically differentiated simple *response suppression*, which refers to simply withholding a pre-potent response, from *attentional control/response conflict*, which encloses the inhibition of an internally represented rule/response set interfering with the ability to engage and implement a new rule/response. This distinction was confirmed in a study with preschoolers (Espy and Bull, 2005) showing that their performance on response conflict, but not on response suppression measures, was related to their simple spans, probably due to

a common reliance on attention control mechanisms. Therefore, the current study set out to investigate the contribution of simple span (verbal or visuospatial) to WM, inhibition (response suppression and verbal and motor response conflict), and shifting in a preschool sample. However, it is important to note that Lerner (2012) failed to find evidence for the proposed dissociation between response suppression and attention control/response conflict in children. A similar less clear-cut distinction between the two inhibitory control dimensions was recently evidenced in a cross-cultural study with preschoolers (Cheie et al., 2014). As an alternative account, Diamond and Kirkham (2005) hypothesized that a common mechanism, called *attentional inertia* (a focus on the same, previously-relevant aspect of one stimulus, even when contextual demands are changing), would be responsible for children's inappropriate responses across various inhibitory and shifting tasks.

Although many tasks have been developed for measuring *shifting* in older children (e.g., Anderson et al., 2000; Jacques and Zelazo, 2001), it is much more difficult to identify comparable tasks for use in the preschool population (Lerner, 2012). In this population, the Dimensional Change Card Sort task (DCCS; Zelazo et al., 2003) has been extensively used to evaluate attention shifting. During task unfolding, children are presented with two target cards (e.g., a blue rabbit and a red boat) and subsequently requested to sort a series of bivalent test cards according to one dimension (e.g., color; the pre-switch phase). After becoming habituated with this dimension, children are asked to sort the same types of test cards according to another dimension (e.g., shape; the post-switch phase). Perseveration on an initial response set shows both the low memory strength of the new mental set (Munakata, 2001), and the reduced ability to inhibit interference from the initial mental set (Diamond et al., 2005). A shifting task either simply involves the coordination of these subordinated skills (Chevalier et al., 2012), or it represents a distinct process acting upon these skills and creating a modification in the original representation of the stimuli (Garon et al., 2008). A modified version of the DCCS was created, using emotional stimuli (facial expressions); the two sorting criteria were emotional expression (happy vs. sad) and gender (Qu and Zelazo, 2007). Children performed significantly better on the emotional faces version (with facilitative effects only in the case of happy faces), suggesting that positive stimuli might promote cognitive flexibility. Since one of our research questions was related to the impact of individual differences in anxiety on EF performance, we constructed an emotional DCCS version (Em-DCCS) similar to this emotional faces version. For this version, we used schematic depictions of facial emotional expressions (sad or happy faces) similar to the ones used by Hadwin and collaborators (2003) in their investigation of anxiety-related biases in visual search. The choice of schematic faces over real emotional expressions was also done in order to eliminate potential cultural effects related to the recognition of facial affect (Posner et al., 1994). The task requires children to switch in the post-shift phase from a neutral judgment (color) to a judgment of emotion (happy or sad faces). Our investigation extends the individual differences direction proposed by Qu and Zelazo (2007) by attempting to replicate the facilitative effect of positive faces on shifting performance, and by relating it to individual differences in trait anxiety. This has been associated with biases in the processing of stimuli with positive versus negative emotional valence in both adults (Chen et al., 2012) and young children (Visu-Petra et al., 2010).

THE ROLE OF INDIVIDUAL DIFFERENCES IN ANXIETY

From an early age, individual differences in anxiety have been shown not only to influence information processing patterns in contexts in which stimuli with emotional valence are present (Pine, 2007; Hadwin and Field, 2010), but also in contexts which lack such emotional information, especially tasks with higher levels of executive demands (see Visu-Petra et al., 2013a, for a review). The explanation of the relationship between individual differences in anxiety and impaired EF has been via the detrimental effects of anxiety on attentional control. This is reflected in the most influential explanatory framework regarding the anxiety—cognitive functioning relationship offered by the Attentional Control Theory (ACT; Eysenck et al., 2007). The theory predicts that in high-anxious individuals, anxiety-related worrisome thoughts interfere with their task-goals, requiring the activation of auxiliary processes and strategies. Accordingly, this concurrent resource activation is mostly evident in decreased performance *efficiency,* as more time and effort are required to complete a task, or to attain a given performance level. Yet, it can also be observed in terms of performance *effectiveness* (response accuracy), especially when the task is more challenging. A compelling body of evidence supports these predictions (see Eysenck and Derakshan, 2011, for a review), confirming that the anxiety-related depletion of resources impedes attention control, diminishing high-anxious individuals' EF (i.e., inhibition, shifting, and updating) performance.

Regarding the impact of anxiety upon preschoolers' STM, predictions are ambivalent. Preliminary evidence shows that, in line with related findings in older children (Hadwin et al., 2005), young children's simple span efficiency and, under certain circumstances, their accuracy, are affected by high trait anxiety levels (Visu-Petra et al., 2009, 2011). Trait anxiety was a longitudinal negative predictor of 3–6 year-old children's verbal STM performance accuracy, as well as efficiency of response, as indicated by a microanalysis of their response time segments (Visu-Petra et al., 2009). Another study revealed that while performance in the visuospatial span tasks did not differ between high-anxious and low-anxious preschoolers, high-anxious 3–7 year-olds displayed an inferior performance on the verbal simple and complex span measures (Visu-Petra et al., 2011). The findings also indicated that on simple span tasks, high-anxious preschoolers displayed efficiency impairments only, while both efficiency and accuracy of response were affected in the verbal WM tasks.

Although the developmental literature directly investigating the effects of anxiety upon EF is scarce (see Visu-Petra et al., 2013b for a review), the existent findings partially support the ACT predictions regarding anxiety's detrimental influence. Specifically, child anxiety has been found to disrupt inhibition efficiency (see Mueller, 2011, for a review), with a cross-cultural study in preschoolers identifying a greater impact of anxiety on performance efficiency in tasks requiring response conflict,

compared to simple response suppression (Cheie et al., 2014). In a context requesting switching between neutral and emotional judgments, higher levels of trait anxiety were found to impair children's performance (Mocan et al., 2014). Several studies have also identified the negative impact of anxiety upon memory updating in younger and older children (e.g., Hadwin et al., 2005; Ng and Lee, 2010; Visu-Petra et al., 2011; Owens et al., 2012). Interestingly, the bidirectional nature of the link between anxiety and EF was recently documented via a longitudinal study that relates EF progress during the transition to school to subsequent teacher ratings of internalizing and externalizing behaviors (Hughes and Ensor, 2011). Additional research is needed to explore how early manifestations of trait anxiety impair attentional control and thus affect executive functioning across neutral or emotionally-salient contexts, and how, in turn, reduced cognitive control further amplifies the information processing patterns specific to anxious cognition and behavior (e.g., Pine, 2007).

CURRENT STUDY

EF dimensions have been shown to undergo intensive developments between the ages of 3 and 6, and their progress during this sensitive developmental window predicted a wide range of cognitive, emotional, and educational outcomes. However, the dependency of these distal outcomes on more basic attentional /memory prerequisites across the preschool years has been theoretically postulated (e.g., Garon et al., 2008), yet not empirically documented. Also, reciprocal links between individual differences in anxiety and various EF dimensions during the preschool years and the transition to school have been identified. However, their interplay has not been systematically investigated. Consequently, the current study was designed to address these two key research questions regarding the developmental EF precursors and early links to individual differences in anxiety, both viewed through the lenses of their early reliance on attentional control mechanisms. Several secondary questions were addressed along the way.

A first aim was to investigate whether EF outcomes (WM, inhibition, shifting) measured during late preschool years could be predicted by children's earlier (9 months) assessed STM spans. We expected greater coherence between measures of verbal and visuospatial WM and their respective STM predictors (Alloway et al., 2006). We also attempted to confirm findings by Espy and Bull (2005), who related measures of response conflict, but not of response suppression, to children's memory spans. To our knowledge, this is the first time that children's performance on a shifting task was related to previous and concurrent levels of STM functioning. A secondary aim was related to the development of STM itself during the preschool years, across the verbal and visuospatial domains. This complements the limited body of longitudinal data documenting intensive progress in children's memory span during this interval (e.g., Gathercole et al., 1992; Schneider et al., 2004; Visu-Petra et al., 2009). The cross-sectional progress for all our measures was followed in order to check for performance improvements in children between 3 and 6 years old.

The second aim concerned the role of individual differences in children's EF performances. In this respect, anxiety-related worrisome thoughts are presumed to generate a cognitive interference, mostly visible in tasks high on executive-demands and/or manipulating verbal information (Eysenck et al., 2007). Hence, we hypothesized that higher levels of anxiety would be related to performance deficits on executive-demanding tasks (especially on verbal WM, response conflict, and set-shifting measures), and to a lesser degree on tasks involving lower executive demands (STM and response suppression). We investigated the role of such individual differences in anxiety while controlling for other individual differences variables such as non-verbal intelligence, age, or gender. Most of our tasks, with the exception of the Em-DCCS, did not require children to process emotional information. Previous studies conducted in the ACT (Eysenck et al., 2007) framework indicate that even in such neutral contexts, especially in high executive-demanding ones, anxiety-related performance deficits can be evident. To our knowledge, this is the first study to systematically link early individual differences in anxiety symptoms to subsequent EF performance.

MATERIALS AND METHODS

PARTICIPANTS AND PROCEDURE

The initial sample consisted of 76 preschoolers recruited from three public kindergartens in the northwest of Romania. However, 8 children could not be followed-up at the second time point (T2), hence data from a total of 68 preschool children (41 boys), aged between 3 years and 2 months and 6 years and 8 months ($M = 4$ years and 8 months, $SD = 10.5$ months) at the first assessment (T1), are presented in the current study. Parents who approved their children's participation were also asked to complete a form requiring demographic information, with exclusion criteria such as neurological or psychological disorders. Aside from parental written consent, the child's verbal assent was also obtained prior to testing. All participants were monolingual Romanian-speaking children, living in urban areas.

Children of parents who gave their written consent were tested individually in a quiet room located at their kindergartens. At T1, all preschoolers were tested in a single session with measures of non-verbal intelligence (Colored Progressive Matrices test), verbal STM (Digit Span and Word Span) and visuospatial STM (Corsi blocks test). Nine months later (at T2), tasks were administered in three separate sessions in order to avoid preschoolers' fatigue and boredom. Hence, in the first session at T2, children were evaluated with the same STM tests administered at T1, with an additional Articulation Rate task, which is not described in this study. Verbal WM (Counting, Backwards Digit, and Listening span) and visuospatial WM tasks (Mr. X, Odd-one-out) were administered in the second session. Finally, inhibition and set-shifting performance were evaluated during a third session (Statue, Knock and Tap, Day/Night Stroop, Em-DCCS), in order to minimize fatigue effects.

MEASURES

Individual differences in intelligence and anxiety

Non-verbal intelligence was assessed using the ***Colored Progressive Matrices test*** (Raven et al., 1998) designed to be suitable for young children. This test consists in 36 individual patterns, for each of which children have to correctly identify the missing

segment (out of 6 possible segments). The total number of correct responses provides a non-verbal intelligence measure for each child.

Trait anxiety was evaluated via parental report on the ***Spence Preschool Anxiety Scale*** (Spence et al., 2001; the Romanian version Benga et al., 2010). The scale consists of 28 anxiety items, 5 non-scored posttraumatic stress disorder items, and another open-ended (non-scored) item. Each parent rated the concordance between the child's behavior and the one described in each item on a 5 point scale. Parents' ratings of the children's anxiety symptoms generated a total score which provided an overall measure of each child's trait anxiety. The trait anxiety measure was administered at T1 only.

Short term memory

During the ***Digit Span*** task (Forward subtest, WISC-III; Wechsler, 1991), children were instructed to repeat each digit sequence spoken by the experimenter in the correct order. The test consists of 9 blocks of 3 trials each. Trials of 2 digits each are included in the first block, after which STM span requirements gradually increase to trials of 9 digits each in the last block. If children correctly recall two trials in a block, the experimenter increases span requirements by moving on to the next block. If the child fails two trials in a block, testing is discontinued.

For the ***Word Span*** task, a list of 9 common two-syllable words was chosen to provide a test of word repetition directly comparable to the other span measures. Two-syllable words were chosen in order to avoid possible word length effects, and to provide a measure more directly comparable to the word length of items from digit span (in Romanian, five out of nine digits have two syllables). Besides stimulus type (words), the task was identical to the Digit Span task in all respects.

Visuospatial STM was evaluated using the ***Corsi blocks test*** (Corsi, 1972). For this test, we used the display provided by the WAIS-R Neuropsychological Inventory (Kaplan et al., 1991). Children were presented with 10 blue cubes randomly located on a board. During task unfolding, the examiner taps a sequence of cubes, and the child is required to reproduce the sequence, by tapping the cubes in the correct order. Besides stimulus type (cube locations), the task was identical to the Digit and Word Span.

STM scoring. Aggregate scores for STM spans were computed following the procedure described by Cowan and collaborators (2003). First, the base span, the highest list length at which the responses for all sequences were correct, was extracted, and a score of 0.33 was added for every correct sequence above this base span. Additionally, a general index of verbal STM was computed by averaging the Word and Digit aggregate spans.

Working memory

WM was evaluated using tasks from the Automated Working Memory Assessment battery (AWMA; Alloway, 2007), a widely-used measure for WM assessment in 4- to 11-year-old children. Three measures were administered in order to assess verbal WM (Counting Recall, Backwards digit recall, Listening recall), while two other (Odd-one-out and Mr. X) were employed to evaluate visuospatial WM. In all these tasks, a particular list length contains 6 trials—if the child correctly performs 3 trials from a list length, the program automatically skips to the next list length. If less than 3 trials from a list length are correctly recalled, testing stops for that task.

In the ***Counting Recall*** test, children are presented with a visual array of red circles and blue triangles. They are asked to count the number of circles in each array, and to memorize the totals. At the end of each trial, children are required to recall the number of circles included in each array, in the correct order. The test consists of 7 blocks of 6 trials each, beginning with trials of one array in the first block, increasing to trials of 7 arrays in the last block.

The ***Backwards Digit Recall*** test is identical to the Digit Span task, except children are required to recall a gradually increasing sequence of spoken digits in the reversed order. The sequences increase by one digit from one block to another, with a maximum of 7 digits for trials corresponding to the last block. The ***Listening Recall*** task consists in a series of short sentences (e.g., "The grass is blue" and "Sugar is sweet") for which children are asked to judge the veracity by giving a "yes" or "no" response to the experimenter. After judging the veracity of each sentence in a trial, children are required to recall the final word of each sentence within the given trial (e.g., "blue" and "sweet"). The test consists of 6 blocks of 6 trials each, with the number of sentences within each trial gradually increasing from two to six.

In the ***Odd-one-out*** task (adapted by the AWMA authors from Russell et al., 1996) children are presented with three shapes, each in a box, displayed in a row. They are then asked to point the odd shape out of each row. After this, the shapes disappear and the child is presented with three empty boxes, being asked to point to where the odd shape was. From the initial level presenting only one row of shapes, difficulty increases up to 7 rows, children being asked to recall the location of the odd shape from each row, in the order they had been shown in each trial. In the ***Mr. X*** task, two fictitious cartoon figures, presented as "Mr. X with the blue hat" and "Mr. X with the yellow hat," are displayed in each item. Children are first asked to identify whether Mr. X with the blue hat is holding *a ball* in the same hand as Mr. X with the yellow hat. With span requirements increasing, more Mr. Xs appear on each block and the child is asked to recall the location of each ball by pointing to a picture with eight compass points. The task consists of 7 blocks of 6 trials each, location span gradually increasing by one with each block.

WM scoring. Aggregate WM spans were computed in the same manner as aggregate STM spans, except that this time a 0.17 score was added to the base span, as one level consisted of 6 trials. Verbal WM and visuospatial WM composite scores were calculated by averaging the scores on corresponding verbal and visuospatial tasks.

Inhibition

In order to assess Inhibition, we used a task requiring simple response suppression, as well as two tasks generating response conflict.

The Statue task from the NEPSY-I battery (Korkman et al., 1998) evaluates *response suppression*, requiring motor persistence

when several distracters are introduced. Children are required to stand in a "statue" position, refraining from vocalizations and body movements for 75-s. During this interval, pre-set distracters are introduced (the examiner coughing, dropping a pen etc.). A 2 points score is attributed for inhibiting any response over each 5-s interval, and a 1 point score for displaying one inappropriate response. The maximum score to be earned by not doing anything throughout this interval is 30.

Knock and Tap is a classical non-verbal Go/No-Go task included in the NEPSY-I battery (Korkman et al., 1998), evaluating *motor response conflict* between immediate motor responses triggered by visual stimuli and the action that is specified in previous verbal directions (Klenberg et al., 2001). In the first part of the test (Part A), children are asked to knock on the table when the examiner taps and vice-versa during 15 trials. In the second part of the task (Part B), children are required to shift to a new set of response. Specifically, they are taught to tap with the side of their first when the examiner knocks and vice-versa, but also to inhibit any motor response when the examiner taps. Part B also consists of 15 trials, and the total number of correct responses (out of 30) determines the accuracy score.

The version of ***Day/Night Stroop*** that we used is an uninterrupted measure of *verbal response conflict*, in which children are presented with a matrix displaying 16 pictures of the sun and moon, respectively. Participants were asked to name the pictures from left to right on each of the four rows, but to inhibit their prepotent responses and say "night" when pointing to the sun, and "day" when pointing to the picture depicting the moon. Thus, we transformed the standard version of the task (Gerstadt et al., 1994) into a more self-paced, speeded task. The maximum accuracy score was 16, and the experimenter timed children's total response in order to obtain an efficiency measure. Accuracy scores may be sufficient for measuring young children's inhibition (e.g., Diamond and Kirkham, 2001), yet in school age children and adults, measures of response time proved to be more sensitive measures, especially when accuracy performance points toward ceiling effects (e.g., MacLeod, 1991; Wright et al., 2003). This later approach was also successfully used with children as young as 3½ years (Simpson and Riggs, 2005). Hence, both latency and accuracy of response were taken into account to generate an inverse-efficiency score (Kennett et al., 2001), calculated as total response time divided by the proportion of correct responses for each participant. Lower values on this measure indicate better inhibitory performance.

Shifting

Finally, shifting performance was estimated using the ***Emotional-Dimensional Change Card Sort (Em-DCCS)***. The classic DCCS task provides a measure of cognitive flexibility in children as young as 3 (Zelazo, 2006). In the emotional version of the task, the target cards consisted of a happy red face and a sad blue face, and their placing (left or right) was counterbalanced across the sample. The version used in this study was modified by using schematic emotional faces, as they were successfully used in previous research regarding anxiety-related bias effects in children (e.g., Hadwin et al., 2003). The schematic faces are presented in **Figure 1**.

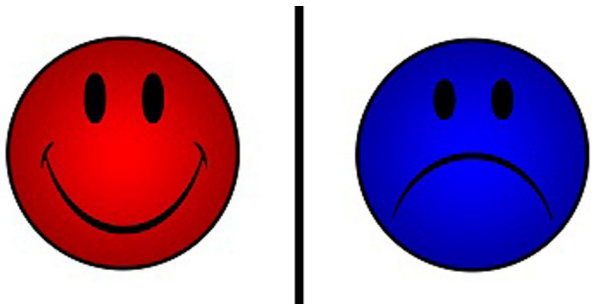


FIGURE 1 | Target cards for the EM-DCCS task.

Participants were initially requested to sort the six pre-switch test cards by the color criterion. After the six pre-switch trials, the experimenter said: "*Now we are going to play another game. We are not going to play the color game anymore. We are going to play the faces game.*" Only performance on the post-switch trials was analyzed, after data from one child who scored poorly (less than 5 out of 6) in the pre-switch phase were excluded. Due to the non-normal distribution of scores on the post-switch phase and the overall high levels of performance, performance was dichotomized using a more stringent criterion than for the pre-switch phase. Thus, two groups were created, children who could perfectly switch to the emotional judgment on all trials of the post-switch phase ($n = 34$) and children with less than perfect performance ($n = 33$).

RESULTS

Analytic approach. In order to determine whether performance on STM tasks was associated with children's EF performance 9 months later, beyond other first assessment measures, separate hierarchical regressions were carried out for each EF outcome (verbal and visuospatial WM, response suppression, verbal and motor response conflict, and attention shifting). The association between individual differences in non-verbal intelligence and trait anxiety was tested in the same manner, after first controlling for the age and gender of the participants. We further tested whether concurrent levels of STM were useful in the prediction of EF outcomes beyond the first assessment STM, age, gender, non-verbal intelligence, and anxiety.

PRELIMINARY ANALYSES

During the univariate and bivariate graphical examination of data, three outlying observations were identified and discarded as they were situated more than 3 SDs below/above the sample means (two on the Day/Night Stroop matrix and one on the Knock and Tap task). One child with poor performance on the pre-switch phase (2/6) of the DCCS task was excluded from the shifting analysis. Univariate descriptives on all measures are listed in **Table 1**, and **Figure 2** presents associations between measures of interest. The correlation matrix for all recorded measures is presented in the Supplementary Materials.

Associations between measures at T1

Older children presented higher non-verbal intelligence (Raven) scores, $r_{(66)} = 0.36$, $p = 0.01$, as well as superior verbal STM,

Table 1 | Univariate descriptives for the raw and composite scores.

Type of measure	Tasks	*N*	Mean	*SD*	Median	Min	Max	Skew	Kurt.
TIME 1									
Age (in months)	–	68	56.00	10.54	56.0	38.00	80.00	0.12	−0.90
Nonvb. intelligence	Raven	68	16.00	4.05	15.50	8.00	27.00	0.59	0.09
Trait anxiety	Spence	68	28.50	14.36	25.0	6.00	68.00	0.55	−0.42
Verbal STM	Word span	66	3.60	0.74	3.70	2.00	5.30	0.02	−0.17
	Digit span	66	4.00	0.84	4.00	2.00	6.30	0.31	0.53
Visuospatial STM	Corsi blocks	68	2.80	0.66	2.70	1.66	5.00	0.81	0.92
TIME 2									
Verbal STM	Word span	68	4.00	0.63	4.00	2.33	5.30	−0.21	−0.28
	Digit span	68	4.20	0.83	4.00	2.33	6.30	0.35	−0.32
Visuospatial STM	Corsi blocks	68	2.90	0.70	2.70	1.33	5.00	0.47	0.17
Verbal WM	Counting span	63	1.62	0.58	1.50	1.00	3.00	0.54	−0.75
	Backward span	63	1.11	0.59	1.00	0.17	2.70	0.19	−0.25
	Listening span	61	1.23	0.67	1.33	0.33	2.30	−0.44	−1.44
Visuospatial WM	Mr. X	63	0.84	0.53	0.84	0.17	1.70	0.27	−1.32
	Odd-one-out	63	1.61	0.51	1.50	0.67	2.70	0.15	−0.77
Response suppression	Statue	68	25.65	4.80	27.00	10.00	30.00	−1.77	2.58
Response conflict	Stroop matrix accuracy	68	13.94	2.97	15.00	2.00	16.00	−1.86	3.77
	Stroop matrix IE	66	1.61	0.77	1.40	0.69	4.20	1.44	2.01
	Knock and Tap	67	27.21	2.78	28.00	18.00	30.00	−1.31	1.46
Shifting	Post-switch DCCS	67	4.24	1.94	6	1	6	−0.33	−1.70
Composite measures	Vb. WM	61	1.32	0.52	1.39	0.50	2.28	0.08	−1.02
	Vs. WM	63	1.23	0.46	1.17	0.42	2.17	0.30	−0.88

Inverse efficiency (IE) was calculated as response time divided by accuracy. Kurt., Kurtosis.

$r_{(64)} = 0.30$, $p = 0.02$, and visuospatial STM spans, $r_{(64)} = 0.49$, $p < 0.001$. Non-verbal intelligence was significantly associated with visuospatial STM, $r_{(64)} = 0.33$, $p < 0.001$, but not with verbal STM, $r_{(64)} = 0.05$, $p < 0.71$. On the other hand, trait anxiety (Spence Preschool Anxiety Scale) negatively correlated with verbal STM span, $r_{(64)} = -0.28$, $p = 0.02$, yet was not associated with visuospatial STM, $r_{(64)} = -0.02$, $p = 0.89$. The results also revealed a significant association between verbal and visuospatial STM composite scales, $r_{(64)} = 0.31$, $p = 0.01$. There were no gender-related differences regarding non-verbal intelligence, anxiety, and STM.

Associations between measures at T2

At T2, there was a again a significant association between verbal and visuospatial STM spans, $r_{(66)} = 0.26$, $p = 0.03$. There was also a positive correlation between verbal and visuospatial WM, $r_{(61)} = 0.51$, $p < 0.001$. A test for the equality of correlations (using the Fisher z transformation) revealed that the correlation between the verbal and visuospatial scales was significantly stronger for WM than for STM, $z = 1.67$, $p = 0.05$, 1-tailed. As expected, verbal STM at T2 correlated positively with verbal WM, $r_{(61)} = 0.62$, $p < 0.001$, while visuospatial STM at T2 was positively associated with visuospatial WM, $r_{(61)} = 0.53$, $p < 0.001$.

The pattern of results regards correlations between WM composite spans and inhibition measures was mixed. Verbal WM correlated positively with motor response conflict (Knock and Tap), $r_{(60)} = 0.41$, $p = 0.01$, and negatively with the (response time based) measure of verbal response conflict (Day/Night Stroop), $r_{(59)} = -0.42$, $p < 0.001$, but did not correlate with response suppression (Statue), $r_{(61)} = 0.18$, $p = 0.17$. Similarly, verbal STM (at T2) correlated positively with motor response conflict (Knock and Tap), $r_{(65)} = 0.38$, $p = 0.01$, and negatively with verbal response conflict (Day/Night Stroop), $r_{(64)} = -0.28$,

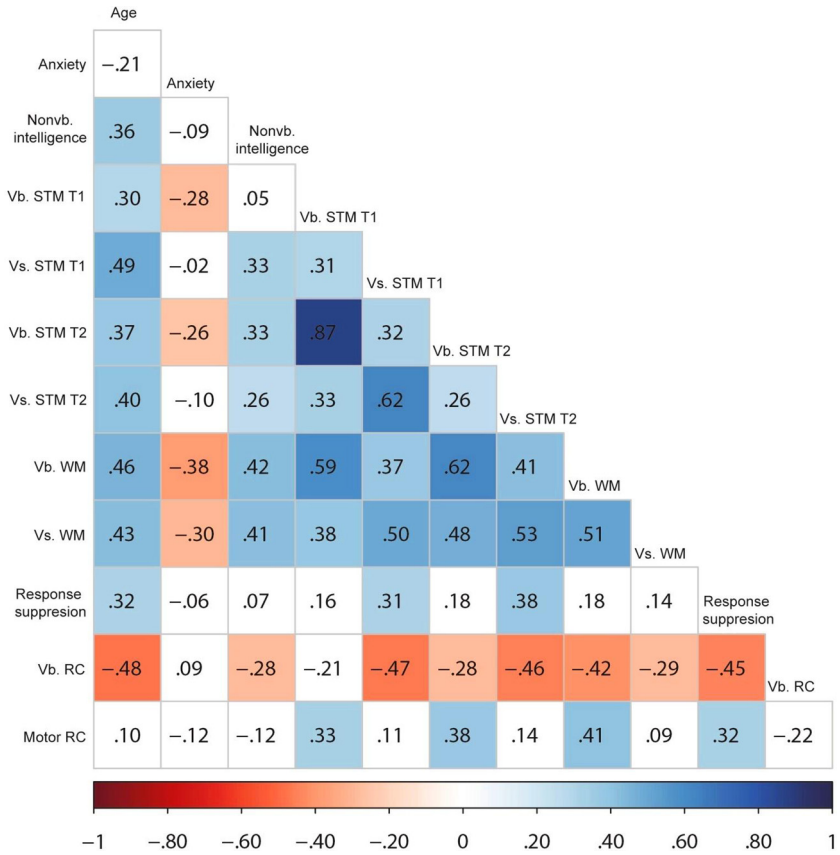


FIGURE 2 | Pearson bivariate correlation matrix for the measures of interest. Color saturation represents the correlation strength according to the scale below the figure. Vb., verbal; Vs., visuospatial; STM, short-term memory; WM, working memory; RC, Response conflict.

$p = 0.02$, but did not correlate with response suppression (Statue), $r_{(66)} = 0.18$, $p = 0.14$. The only inhibition measure associated with visuospatial WM was verbal response conflict (Day/Night Stroop), $r_{(59)} = -0.29$, $p = 0.02$. Visuospatial STM (at T2) correlated significantly with both verbal response conflict (Day/Night Stroop), $r_{(64)} = -0.46$, $p = 0.01$, and response suppression (Statue), $r_{(66)} = 0.38$, $p = 0.01$. The correlation between motor (Knock and Tap) and verbal (Day/Night Stroop) response conflict was non-significant, $r_{(64)} = -0.22$, $p = 0.07$. However, response suppression (Statue) correlated with both motor, $r_{(65)} = 0.32$, $p = 0.01$, and verbal response conflict, $r_{(64)} = -0.45$, $p < 0.001$.

Longitudinal associations

The associations between STM spans at the two time points were substantial, particularly for verbal tasks, $r_{(64)} = 0.87$, $p < 0.001$. The gains in STM (calculated as the difference between T2 and T1 spans) correlated significantly and negatively with the corresponding STM spans at T1, $r_{(64)} = -0.48$, $p < 0.001$, for verbal STM, and $r_{(66)} = -0.37$, $p = 0.01$, for visuospatial STM. A paired t-test revealed that gains in verbal STM were highly significant, as the difference between the two time points was, on average 0.26 (95% CI from 0.17 to 0.35). The visuospatial STM gains were also significant, with a mean difference of 0.14 (95% CI from 0.01 to 0.29), but the estimate of the mean difference lacked precision due to the large variance in gains ($SD = 0.59$). The results revealed no significant links between the STM gains and anxiety, or non-verbal intelligence. STM spans at T1 correlated moderately with the corresponding WM spans, $r_{(59)} = 0.59$, $p < 0.001$, for verbal measures, and $r_{(61)} = 0.50$, $p < 0.001$, for visuospatial measures.

With regards to associations with the individual differences measured at T1, results revealed that non-verbal intelligence was positively associated to verbal WM scores, $r_{(61)} = 0.42$, $p < 0.001$, and visuospatial WM, $r_{(61)} = 0.41$, $p < 0.001$. The only other EF measure associated with non-verbal intelligence was the Day/Night Stroop inverse efficiency, $r_{(64)} = -0.28$, $p = 0.02$, revealing that children with higher non-verbal intelligence scores also had superior performances in terms of verbal response conflict (Day/Night Stroop). At the same time, correlations also revealed that higher anxiety was linked to lower verbal STM spans at T2, $r_{(66)} = -0.26$, $p = 0.04$, as well as to lower verbal and visuospatial WM spans, $-0.38 < r < -0.30$. However, trait anxiety was not significantly related to response conflict (Knock and Tap, and Day/Night Stroop) or response suppression (Statue). The mean anxiety score of children who did not pass the shifting

task (DCCS, $M = 32.20$, $SD = 12.10$) was significantly higher than that of the children who passed ($M = 24.79$, $SD = 15.85$), $t_{(60)} = 2.17$, $p = 0.03$. The only T2 measure for which gender effects were found was attention shifting as the odds of maximal performance for girls were 6.11 times (95% CI from 1.99 to 18.76) the odds of boys.

CROSS-SECTIONAL EFFECTS OF AGE

The current section charts the age-related progress in both STM and EF abilities through a descriptive, cross-sectional approach. The graphical exploration in **Figure 3A** suggests that the most substantial improvements in terms of verbal STM span roughly occurred between the age of 3–4 to 4–5 years, after which performance stagnated or had a more modest increase up to the age of 6–7 years. The only significant increase in verbal STM performance at T1 was evident when comparing 3- ($M = 3.25$, $SD = 0.74$) to 4-year-olds ($M = 4.05$, $SD = 0.62$), $t_{(30)} = 3.72$, $p < 0.001$. However, at T2, 5-year-olds ($M = 4.23$, $SD = 0.62$) significantly outperformed 4-year-olds ($M = 3.69$, $SD = 0.64$), $t_{(32)} = 2.83$, $p = 0.01$. This discrepancy made it difficult to pinpoint the exact age at which peak performance in verbal STM was achieved. However, it is certain that 6-year-olds did not outperform 5-year-olds in terms of verbal STM at any time point.

For visuospatial STM (see **Figure 3A**), children's improvement was more gradual and continuous across the whole age range. Based on the T2 assessment of their visuospatial STM span performance, the difference in estimated means between children 1-year apart was of 0.32 (95% CI from 0.16 to 0.48).

The improvements in WM spans (see **Figure 3B**) followed a fairly linear trend, although there was considerable variability in performance within each 1 year age range ($SDs = 0.40$). Older children had mean verbal WM spans 0.27 higher (95% CI from 0.14 to 0.40) than their 1 year younger peers. Older children also had a 0.23 (95% CI from 0.11 to 0.34) increase in estimated mean visuospatial WM span.

Similarly, the response time-based measure of verbal response conflict (i.e., log Day/Night Stroop inverse efficiency) followed a downward linear trend as age increased (see **Figure 3C**). A one-year increase in age was related to a change in estimated mean Day/Night Stroop performance of −0.39 (95% CI from −0.59 to −0.20). However, older children did not outperform their younger peers on motor response conflict (Knock and Tap) as a 1-year increase in age was associated with a 0.13 increase in mean motor response conflict (95% CI from −0.25 to 0.52). On the response suppression task, a 1-year increase in age was associated with an increase of 1.73 (95% CI from 0.48 to 2.98) in estimated mean response suppression (Statue). The percentage of children passing on the post-shift DCCS increased from 25% (for 4-year olds) to 67% (for 6-year olds).

CONCURRENT AND PREDICTIVE EFFECTS

Separate eight step hierarchical regressions were conducted for each outcome (verbal and visuospatial WM, response

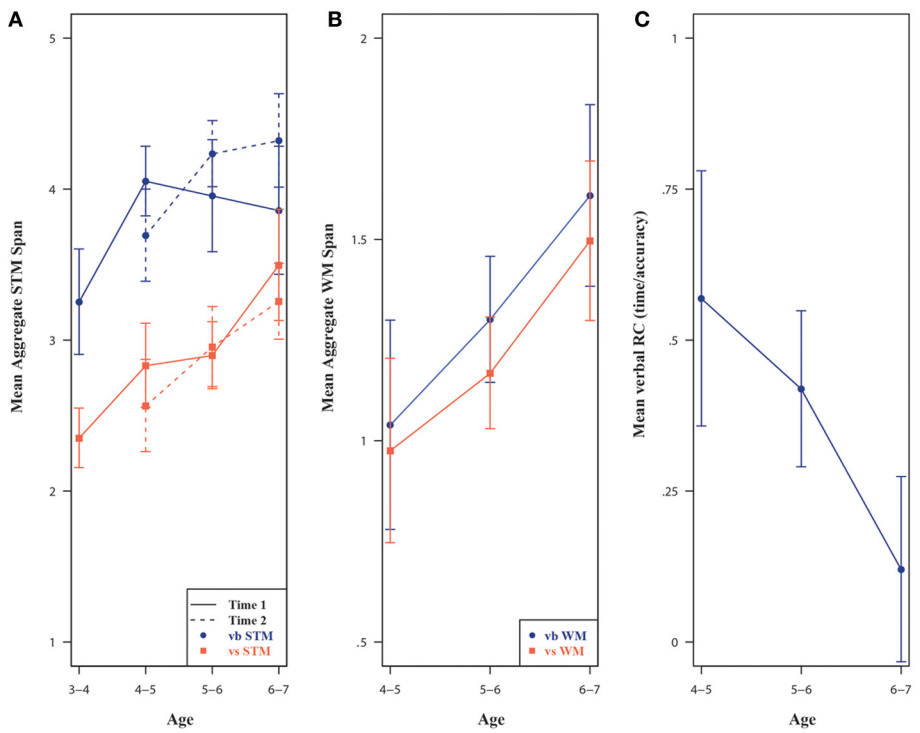


FIGURE 3 | Children's mean STM (A), WM (B), and verbal response conflict (C) performances with 95% Confidence Intervals. Performance for verbal response conflict was assessed using the log inverse efficiency score (total time to complete the task/total score). Vb., verbal; Vs., visuospatial; STM, short-term memory; WM, working memory; RC, Response conflict.

suppression, verbal and motor response conflict, and shifting). For all outcomes, the first four predictors included in the regressions were: age at T1 (step 1), gender (step 2), non-verbal intelligence (step 3), and trait anxiety (step 4). In step 5, verbal STM at T1 was added as a predictor of verbal WM, and visuospatial STM at T1 as a predictor of visuospatial WM. In the subsequent steps, we added the domain non-specific STM measures at T1 (i.e., verbal STM for visuospatial WM, and visuospatial STM for verbal WM; step 6), followed by the domain specific STM measures at T2 (step 7) and the domain non-specific STM measures at T2 (step 8). For all other EF outcomes, the remaining predictors were added as follows: verbal STM at T1 (step 5), visuospatial STM at T1 (step 6), verbal STM at T2 (step 7), visuospatial STM at T2 (step 8). Hierarchical regression models with the coefficient of determination (R^2) at each step and the F-tests comparing consecutive models are presented in **Table 2** for WM, as well as in **Table 3** for inhibition and shifting.

Working memory

Individual differences in age, non-verbal intelligence, and anxiety differently accounted for children's WM variance. Accordingly, the first predictor considered, age, accounted for the largest proportion of variance explained in both verbal and visuospatial WM, while the addition of gender in the second step did not benefit the models. Non-verbal intelligence improved significantly only the model of verbal WM span. Further, individual differences in anxiety were associated to significant changes in the amount of variance explained in both WM spans, beyond the contributions of children's age and non-verbal intelligence scores. Overall, domain-specific STM measured at T1 was a very good predictor of the respective domain-specific WM, explaining as much as 18% of variance in verbal WM performance, after considering the effects of age, gender, non-verbal intelligence, and trait anxiety. On the other hand, domain non-specific STM did not improve either WM model. Controlling for previous (T1) STM spans, the addition of concurrent domain specific (T2) STM measures did not significantly improve the verbal WM model, but had a small significant effect on visuospatial WM. However, multicollinearity (*VIF*s as high as 5.6) between the STM spans at the two time points made it difficult to make inferences about individual (STM at T1 or at T2) predictors. We were primarily interested in the ability to predict WM based on STM at T1, therefore, we relied on the models in the sixth step of the hierarchical regressions to quantify this relationship (see **Table 2**).

In the final verbal WM model, the best predictor was verbal STM; a one point increase in verbal STM was associated with a change of 0.31 (95% CI from 0.17 to 0.44) in the estimated mean verbal WM span (see **Figure 4A**) keeping all other predictors constant. Also, higher non-verbal intelligence scores were linked to higher verbal WM performance, $b = 0.04$, $SE = 0.01$, $p = 0.01$. Children with higher anxiety scores tended to have lower verbal WM spans, $b = -0.007$, $SE = 0.003$, $p = 0.05$. This result is illustrated in **Figure 4C**. Lastly, age, gender, and visuospatial STM (T1) were not significant in the final model.

Similarly, in the final model for visuospatial WM performance (step 6), visuospatial STM was the only significant predictor; a one point increase in STM span was associated with a 0.22

Table 2 | Hierarchical regression models predicting children's performance on verbal and visuospatial working memory (WM) spans.

	Verbal WM			Visual WM		
	B	***SE***	***β***	***B***	***SE***	***β***
STEP 1						
Intercept	0.03	0.33	–	0.15	0.29	–
Age	0.02	0.01	0.45***	0.02	0.00	0.43***
$R^2(\Delta F)$	0.21 (15.31***)			0.19 (13.50***)		
STEP 2						
Intercept	−0.004	0.33		0.15	0.30	
Age	0.02	0.01	0.32***	0.02	0.00	0.43***
Gender	0.17	0.12	0.16	−0.01	0.11	−0.01
$R^2(\Delta F)$	0.23(1.96)			0.19 (0.01)		
STEP 3						
Intercept	−0.26	0.34	–	−0.02	0.31	–
Age	0.02	0.01	0.32*	0.01	0.01	0.34*
Gender	0.20	0.12	0.19	0.01	0.11	0.01
Nonvb. Intelligence	0.04	0.02	0.28*	0.03	0.02	0.22
$R^2(\Delta F)$	0.30 (5.32*)			0.23 (2.91)		
STEP 4						
Intercept	0.32	0.38	–	0.38	0.36	–
Age	0.01	0.01	0.24*	0.01	0.01	0.28*
Gender	0.21	0.11	0.19	0.01	0.10	0.01
Nonvb. Intelligence	0.03	0.02	0.25	0.02	0.01	0.20
Anxiety	−0.01	0.00	−0.31**	−0.01	0.00	−0.25*
$R^2(\Delta F)$	0.39 (8.08**)			0.28 (4.44*)		
STEP 5						
Intercept	−0.83	0.40	–	0.15	0.35	–
Age	0.01	0.00	0.16	0.00	0.01	0.11
Gender	0.17	0.09	0.16	0.00	0.01	0.00
Nonvb. Intelligence	0.04	0.01	0.28**	0.02	0.01	0.17
Anxiety	−0.01	0.00	−0.19$^+$	−0.01	0.00	−0.26*
Domain specific STM (T1)	0.32	0.07	0.46***	0.25	0.09	0.35**
$R^2(\Delta F)$	0.57 (23.18***)			0.37 (7.98**)		
STEP 6						
Intercept	−0.86	0.40	–	−0.25	0.41	–
Age	0.01	0.01	0.12	0.00	0.01	0.09
Gender	0.17	0.09	0.16	−0.01	0.01	−0.01
Nonvb. Intelligence	0.04	0.01	0.27**	0.02	0.01	0.19
Anxiety	−0.01	0.00	−0.19*	−0.01	0.00	−0.20
Domain specific STM (T1)	0.31	0.07	0.44***	0.22	0.09	0.30*
Domain non-specific STM (T1)	0.07	0.09	0.09	0.12	0.07	0.20
$R^2(\Delta F)$	0.57 (0.69)			0.41 (2.99)		
STEP 7						
Intercept	−0.95	0.44	–	−0.28	0.40	–
Age	0.00	0.01	0.11	0.00	0.01	0.09
Gender	0.17	0.10	0.16	−0.01	0.09	−0.01
Nonvb. Intelligence	0.03	0.01	0.26*	0.02	0.01	0.15
Anxiety	−0.01	0.00	−0.18	−0.01	0.00	−0.20
Domain specific STM (T1)	0.24	0.14	0.34	0.10	0.10	0.14
Domain non-specific STM (T1)	0.08	0.09	0.10	0.09	0.07	0.15

(Continued)

Table 2 | Continued

	Verbal WM			Visual WM		
	B	***SE***	***β***	***B***	***SE***	***β***
Domain specific STM (T2)	0.09	0.17	0.11	0.18	0.08	0.30*
$R^2(\Delta F)$	0.58 (0.29)			0.46 (4.89**)		
STEP 8						
Intercept	−0.99	0.44	–	−0.48	0.43	–
Age	0.00	0.01	0.10	0.00	0.01	0.06
Gender	0.18	0.10	0.17	0.01	0.01	0.01
Nonvb. Intelligence	0.03	0.01	0.24*	0.01	0.01	0.10
Anxiety	−0.01	0.00	−0.18	−0.00	0.00	−0.18
Domain specific STM (T1)	0.21	0.14	0.30	0.11	0.10	0.16
Domain non-specific STM (T1)	0.03	0.10	0.04	−0.06	0.14	−0.10
Domain specific STM (T2)	0.12	0.17	0.14	0.20	0.08	0.32*
Domain non-specific STM (T2)	0.08	0.09	0.12	0.20	0.17	0.29
$R^2(\Delta F)$	0.58 (0.98)			0.47 (1.46)		

β, Standardized regression coefficient; T1, first time assessment; T2, second time assessment; Domain specific STM, verbal STM for verbal WM, and visuospatial STM for visuospatial WM. Domain non-specific STM, visuospatial STM for verbal WM, and verbal STM for visuospatial WM. The baseline gender is male. $^{+}p < 0.06$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

increase (95% CI from 0.04 to 0.40) in estimated mean WM span (see **Figure 4B**), given the other model predictors. The age, gender, non-verbal intelligence, trait anxiety levels, and verbal STM (T1) did not prove significant. According to the bootstrapped R^2, denoting the ratio of explained variance, the verbal WM model (bootstrapped $R^2 = 0.50$, 95% BCa CI from 0.35 to 0.68) performed relatively well. However, the visuospatial WM model did not match this performance, bootstrapped $R^2 = 0.32$ (95% BCa CI from 0.19 to 0.52).

Inhibition

Response suppression. The hierarchical regression revealed that beyond the first step (age), no other variable improved the model for response suppression (see **Table 3**). This could be explained by a lack of variability in the outcome measure, nearing a ceiling effect. The effect of age remained significant in the model including gender, non-verbal intelligence, anxiety and verbal STM (T1), but was insignificant in subsequent models. A description of the relationship between age and response suppression is presented in section Cross-sectional effects of age.

Response conflict. The first significant improvement in the model for *motor response conflict* (Knock and Tap) came with the addition of verbal STM at T1 in the fifth step, although the addition of non-verbal intelligence (step 3) was marginally significant. In the final model (step 6), only verbal STM remained significant when controlling for age, gender, non-verbal intelligence, anxiety, and visuospatial STM at T1, $b = 1.07$, $SE = 0.49$, $p = 0.030$. The scarcity of good predictors is probably related to the fact that, on this task, the performance of the majority of children was very good and there was little variability in the outcome measure. The overall model (step 6) performed poorly, bootstrapped $R^2 = 0.12$ (95% BCa CI from 0.05 to 0.26).

For *verbal response conflict*, the addition of any variables after the first step (age) proved inconsequential in improving the model fit. Age continued to be a good predictor in the model containing age, gender, non-verbal intelligence, and verbal and visuospatial STM at T1 (step 6), $b = -0.01$, $SE = 0.01$, $p = 0.030$. Despite this association, the verbal response conflict model performed less well overall as compared to both WM models, bootstrapped $R^2 = 0.21$ (95% BCa CI from 0.11 to 0.40).

Shifting

The probability of passing the DCCS test provided a measure of children's shifting performance. The hierarchical regression (see **Table 3**) revealed that besides age, the addition of gender, anxiety, and verbal STM at T1 improved previous models. In the final model (step 6), age was no longer significant, alongside non-verbal intelligence, trait anxiety and visuospatial STM at T1. However, keeping all else constant, verbal STM span at T1 was a useful predictor of shifting performance. The estimated probability of success was larger for children with better verbal STM performance as the odds of success were 2.99 times (95% CI from 1.06 to 8.40) larger for children who had verbal STM spans larger by one unit than their peers. The DCCS was the only measure for which we observed gender differences. The odds ratio of success for the girls relative to the boys was 11.74 (95% CI from 2.70 to 51.04), given the same age, non-verbal intelligence, anxiety, and STM spans. A graphical representation of the predicted probabilities of success as a function of gender and verbal STM at T1 is provided in **Figure 5**. The performance of the model in terms of (Cox and Snell's[1]) R^2 was comparable to the one of the WM models, bootstrapped $R^2 = 0.36$ (95% BCa CI from 0.24 to 0.53).

We were also interested in whether there were differences in the post-switch DCCS performance between sad and happy stimuli. A McNemar's test failed to show any differences related to the emotionality of the faces, $\chi^2(1) = 0.55$, $p = 0.46$. Further, we wanted to explore whether anxiety influenced DCCS trials with different expressions to a similar extent using a more sensitive measure of performance based on accuracy, rather than a pass/fail criterion. Two Poisson regressions were carried out for each expression including gender, anxiety, and verbal STM as predictors. Anxiety was a significant predictor for sad post-switch trials, $b = -0.014$, $SE = 0.007$, $p = 0.05$, but not for happy post-switch trials, $b = -0.005$, $SE = 0.006$, $p = 0.45$.

DISCUSSION

Our study addressed two major research questions. The first one was developmental, concerning the interrelationships between early levels of STM performance and subsequent levels of the same ability, assessing their predictive value for three EF dimensions: WM, inhibition, and shifting. The second one was an

[1]Note that the maximum of the Cox and Snell pseudo R^2 *is lower than 1* $(1 - L(M_{Intercept})^{2/N})$.

Table 3 | Hierarchical regression models predicting children's performance on inhibition (verbal and motor response conflict, and response suppression) and shifting measures.

	Verbal RC			Motor RC			Response suppression			Shifting		
	B	***SE***	***β***	***B***	***SE***	***β***	***B***	***SE***	***β***	***B***	***SE***	***β***
STEP 1												
Intercept	1.52	0.26		25.64	1.90		17.33	3.07		−3.67	1.50	
Age	−0.02	0.00	−0.49***	0.03	0.03	0.10	0.15	0.05	0.33**	0.07	0.03	0.68*
$R^2(\Delta F)$		0.24 (19.80***)			0.01 (0.66)			0.11 (7.64**)			0.10 (6.80**)	
STEP 2												
Intercept	1.52	0.26		25.58	1.92		17.20	3.10		−2.83	1.64	
Age	−0.02	0.00	−0.49***	0.03	0.03	0.10	0.15	0.05	0.32**	0.07	0.03	0.77*
Gender	0.01	0.10	0.02	0.17	0.72	0.03	0.51	1.17	0.05	−2.02	0.63	−2.02***
$R^2(\Delta F)$		0.24 (0.02)			0.01 (0.06)			0.11 (0.19)			0.25 (12.15***)	
STEP 3												
Intercept	1.62	0.29		27.20	2.06		17.65	3.43		−4.17	1.91	
Age	−0.02	0.00	−0.46***	0.05	0.03	0.20	0.15	0.06	0.34*	0.06	0.03	0.59
Gender	0.01	0.10	0.01	0.05	0.70	0.01	0.47	1.19	0.05	−2.23	0.66	−2.23***
Nonvb. intelligence	−0.01	0.01	−0.10	−0.19	0.10	−0.26	−0.05	0.17	−0.04	0.15	0.09	0.62
$R^2(\Delta F)$		0.25 (0.73)			0.07 (3.81$^+$)			0.11 (0.10)			0.29 (3.13)	
STEP 4												
Intercept	1.62	0.34		28.89	2.36		17.49	4.05		−2.18	2.13	
Age	−0.02	0.00	−0.46***	0.05	0.03	0.17	0.16	0.06	0.34*	0.05	0.03	0.56
Gender	0.01	0.10	0.01	0.07	0.70	0.01	0.47	1.20	0.05	−2.43	0.71	−2.43***
Nonvb. intelligence	−0.01	0.01	−0.10	−0.21	0.10	−0.29*	−0.05	0.17	−0.04	0.14	0.09	0.55
Anxiety	0.00	0.00	−0.00	−0.04	0.02	−0.18	0.00	0.04	0.01	−0.05	0.02	−0.78*
$R^2(\Delta F)$		0.25 (0.00)			0.10 (2.05)			0.11 (0.01)			0.34 (5.25*)	
STEP 5												
Intercept	1.72	0.41		25.36	2.78		16.01	4.98		−6.04	3.01	
Age	−0.02	0.00	−0.44***	0.02	0.04	0.09	0.15	0.06	0.32*	0.03	0.04	0.35
Gender	0.01	0.10	0.01	−0.08	0.68	−0.01	0.42	1.21	0.04	−2.47	0.75	−2.47***
Nonvb. intelligence	−0.01	0.01	−0.11	−0.19	0.10	−0.26	−0.04	0.17	−0.03	0.18	0.10	0.72
Anxiety	0.00	0.00	−0.02	−0.02	0.02	−0.11	0.01	0.04	0.03	−0.05	0.03	−0.66
Vb. STM (T1)	−0.03	0.07	−0.06	1.07	0.48	0.29*	0.46	0.85	0.07	1.08	0.51	0.81
$R^2(\Delta F)$		0.25 (0.21)			0.17 (4.94*)			0.11 (0.27)			0.39 (5.25*)	
STEP 6												
Intercept	1.77	0.40		25.37	2.81		15.56	4.95		−6.08	3.01	
Age	−0.01	0.01	−0.32*	0.02	0.04	0.10	0.10	0.07	0.23	0.03	0.04	0.36
Gender	0.01	0.10	0.01	−0.08	0.68	−0.01	0.41	1.20	0.04	−2.46	0.75	−2.46***
Nonvb. intelligence	−0.01	0.01	−0.08	−0.19	0.10	−0.26	−0.06	0.17	−0.05	0.18	0.10	0.72
Anxiety	0.00	0.00	0.01	−0.02	0.03	−0.11	0.00	0.00	0.01	−0.05	0.03	−0.65
Vb. STM (T1)	−0.01	0.07	−0.02	1.07	0.49	0.29*	0.12	0.87	0.03	1.09	0.53	0.82*
Vs. STM (T1)	−0.17	0.09	−0.25	−0.05	0.65	−0.01	1.53	1.13	0.20	−0.04	0.61	−0.03
$R^2(\Delta F)$		0.30 (3.40)			0.17 (0.01)			0.14 (1.85)			0.39 (0.01)	
STEP 7												
Intercept	1.90	0.42		23.47	2.92		13.95	5.27		−5.32	3.09	
Age	−0.01	0.01	−0.30*	0.01	0.04	0.05	0.09	0.07	0.21	0.06	0.04	0.57
Gender	0.01	0.10	0.00	0.04	0.67	0.01	0.51	1.21	0.05	−2.62	0.80	−2.62***
Nonvb. intelligence	−0.01	0.01	−0.05	−0.25	0.10	−0.34*	−0.11	0.18	−0.09	0.26	0.12	1.07*
Anxiety	0.00	0.00	0.00	−0.01	0.03	−0.07	0.01	0.04	0.03	−0.06	0.03	−0.85
Vb. STM (T1)	0.09	0.14	0.15	−0.56	0.98	−0.15	−1.18	1.76	−0.18	2.90	1.19	2.17*
Vs. STM (T1)	−0.18	0.09	−0.27$^+$	0.14	0.64	0.03	1.69	1.15	0.22	−0.24	0.68	−0.16
Vb. STM (T2)	−0.14	0.17	−0.21	2.20	1.14	0.52$^+$	1.86	2.07	0.25	−2.22	1.28	−1.53
$R^2(\Delta F)$		0.30 (0.66)			0.22 (3.71$^+$)			0.15 (0.81)			0.42 (3.38)	

(Continued)

Table 3 | Continued

	Verbal RC			Motor RC			Response suppression			Shifting		
	B	***SE***	***β***	***B***	***SE***	***β***	***B***	***SE***	***β***	***B***	***SE***	***β***
STEP 8												
Intercept	1.96	0.42		23.19	2.93		13.05	5.17		−5.12	3.11	
Age	−0.01	0.01	−0.30*	0.01	0.04	0.04	0.09	0.07	0.19	0.05	0.04	0.58
Gender	−0.01	0.10	−0.01	0.07	0.67	0.01	0.59	1.18	0.06	−2.66	0.81	−2.66***
Nonvb. intelligence	0.00	0.01	−0.01	−0.27	0.10	−0.36*	−0.18	0.18	−0.14	0.27	0.13	1.11*
Anxiety	0.00	0.00	0.00	−0.01	0.02	−0.07	0.01	0.04	0.03	−0.06	0.03	−0.85
Vb. STM (T1)	0.15	0.14	0.26	−0.80	1.00	−0.22	−1.93	1.76	−0.30	2.97	1.21	2.23*
Vs. STM (T1)	−0.08	0.11	−0.11	−0.24	0.74	−0.05	0.41	1.30	0.05	−0.06	0.81	−0.04
Vb. STM (T2)	−0.18	0.16	−0.27	2.38	1.16	0.56*	2.44	2.04	0.33	−2.29	1.30	−1.58
Vs. STM (T2)	−0.17	0.09	−0.28	0.65	0.63	0.16	2.14	1.10	0.31	−0.28	0.66	−0.20
$R^2(\Delta F)$	0.35 (3.54)			0.23 (1.06)			0.21 (3.78)			0.43 (0.18)		

β, Standardized regression coefficient. For shifting, β was obtained by standardizing the continuous predictors, and Cox and Snell's R^2 was calculated. T1, first time assessment; T2, second time assessment. Vb., verbal; Vs., visuospatial, RC, response conflict. The baseline gender is male. $^{+}p < 0.06$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

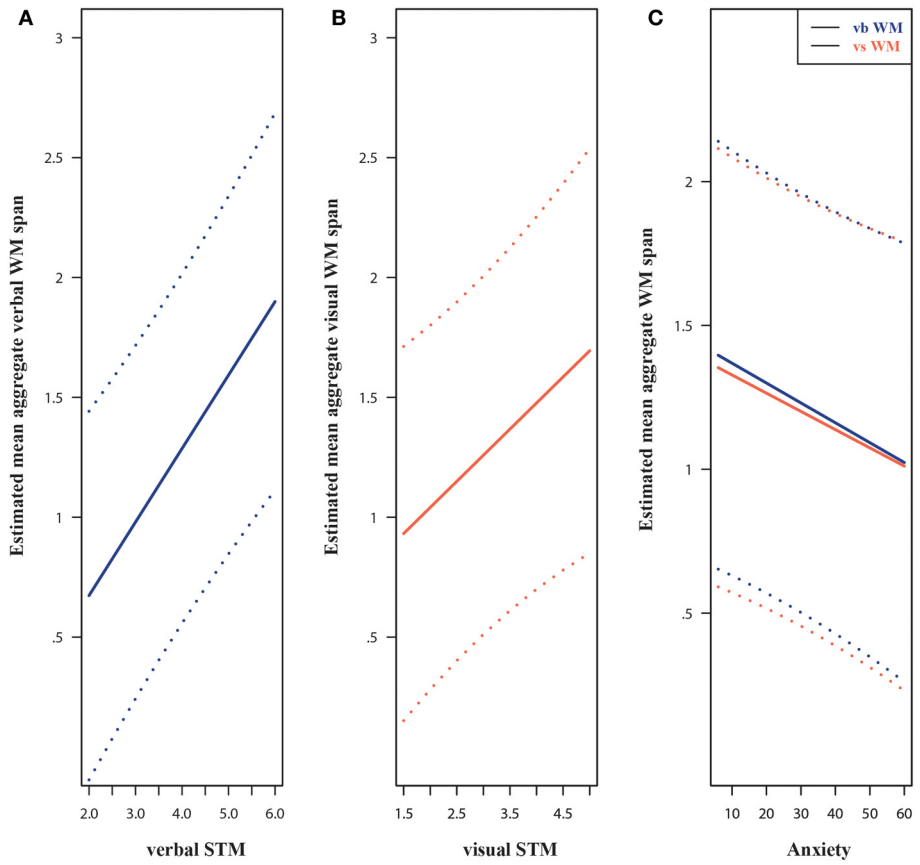


FIGURE 4 | Estimated regression lines and 95% Prediction Intervals for: verbal STM and verbal WM (A), visuospatial STM and visuospatial WM (B), and for trait anxiety and WM (C). Estimated means correspond to boys and all other (non-significant) continuous model predictors were set to the mean sample values. Vb., verbal; Vs., visuospatial; STM, short-term memory; WM, working memory.

individual differences question, and it concerned the predictive value of early levels of anxiety on subsequent EF, controlling for the influence of other individual differences in age, gender, or non-verbal intelligence. In the literature, both anxiety and STM, have been linked to attention control (dys)functions, making it plausible to assume that attention control could represent a mechanism responsible for their association with EF performance.

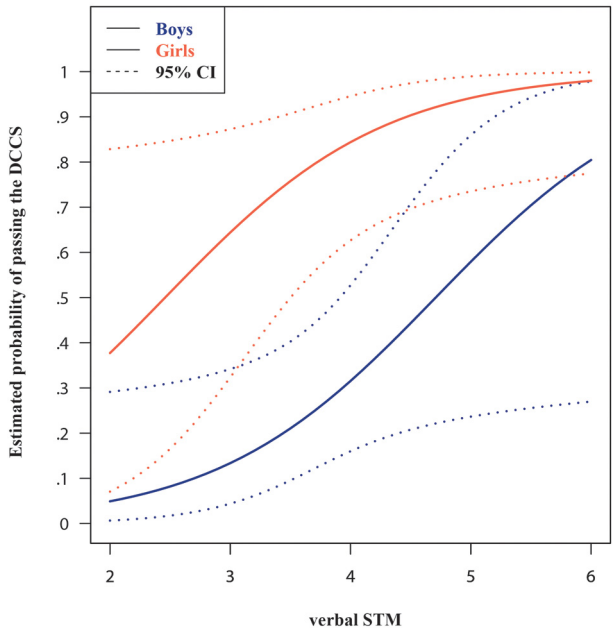


FIGURE 5 | Estimated probabilities of success on the shifting (DCCS) task, as a function of verbal STM span and gender. Estimated means correspond to boys and all other (non-significant) continuous model predictors were set to the mean sample values. STM, short-term memory.

EARLY EF DEVELOPMENT AND ITS PRECURSORS

A preliminary analysis of inter–task *correlations* revealed stronger relationships between measures designed to tap the same underlying memory component, confirming domain specificity. The verbal and visuospatial scales correlated better for WM than for STM, confirming that WM measures using different stimuli actually capture more common underlying cognitive processes than STM tasks. This fits nicely with the suggestion that controlled attention works to keep task-relevant information active in WM across a variety of stimulus modalities (Engle et al., 1999). Interestingly, reanalyzing correlations among the three inhibitory control measures, we noted that the proposed dissociation between response suppression and response conflict measures (e.g., Friedman and Miyake, 2004; Espy and Bull, 2005) was not fully validated. More specifically, while the associations between the verbal and motor response conflict were poor, scores on response suppression were significantly related to verbal response conflict. The lack of a significant association between the two response conflict measures could have different explanations, including the different outcome measures (accuracy vs. inverse efficiency), the use of different stimuli (verbal vs. visuospatial), or a truly modest coherence between various inhibitory control measures in young children (see Lerner, 2012; Van der Ven et al., 2012; Cheie et al., 2014). However, although expected for this age range (e.g., Willoughby et al., 2012b), the high levels of performance reached in most inhibitory task preclude us from drawing strong conclusions regarding the independence or interdependence of various inhibitory control measures.

The cross-sectional analysis of the evolution of STM and EF abilities within this developmental period revealed different growth patterns for the various outcomes measured. The mean *verbal STM span* improved over the course of 3 years by roughly one unit, meaning that while the youngest children (aged 3) had a mean maximal span of 3, the oldest ones (aged 6) had a mean maximal span of 4. Verbal STM performance during this period most likely reached its peak sometime between the ages of 4 and 5 years, and then in the transition to 6 years progress stagnated. This confirms previous research indicating that performance in verbal STM tasks levels off sooner than in visuospatial ones, although the exact level at which this plateau occurs is placed later, at about 10–11 years (Alloway et al., 2006). This suggests that our findings might indicate simply a transitory slowing down of verbal STM progress in the late preschool years. However, this does not imply that there is no within-individual gain in verbal STM, as such gains were evident in our study, over the 9 month period, and did not vary as a function of age. Moreover, it is plausible that over this apparent stagnation period, lower performing children may still continue to improve so as to match their peers. This statement can be supported by the negative correlation between STM at the first measurement point and the within-participant gain in STM.

The *development of visuospatial STM* was more gradual, performance increasing linearly within the age range recorded in the study (between 3 and 6 years of age). This parallels previous proofs of a steady increase in performance on tests that employ visual material that is not phonologically recordable (e.g., Pickering et al., 2001). A storage hypothesis has been proposed (Logie and Pearson, 1997), suggesting the increase in the capacity of the visuospatial sketchpad. Alternatively, an increasing involvement of the central executive has been suggested via more effective strategies or long-term memory knowledge deployment (Pickering et al., 2001). However, a better rate of attention shifting between locations could also be responsible for the increase in spatial span (Smith and Jonides, 1997). Using a similar measure across tasks (the aggregate span) allows us to directly contrast absolute levels of performance on the verbal compared to the visuospatial STM. Evident from the descriptive analyses, visuospatial STM in the oldest children (6–7 years) did not match comparable levels of verbal STM at a much younger age (4–5 years). This confirms the well-documented inferior visuospatial span compared to the verbal one in preschool children (e.g., Pickering et al., 1998; Alloway et al., 2006). The fact that these tasks are experienced as more difficult is consequential for their greater involvement in some EF tasks discussed below, confirming that visuospatial STM tasks draw more executive resources than the verbal STM measures (Miyake et al., 2001; Alloway et al., 2006).

Age-related improvements regarding children's *WM performance* appeared to be more gradual, similar to previously identified trends (Alloway et al., 2006). The mean aggregate span increased with roughly half a unit over the course of three years on both verbal and visuospatial WM measures. During the developmental course of WM, it appears that domain-general processing mechanisms interact with domain-specific storage components leading to a gradual progress (Bayliss et al., 2003, 2005; Alloway

et al., 2006) also documented in the current study. Modest age-related improvements in performance also occurred on the *motor response conflict* and the *response suppression* task. However, there were no age-related improvements on the accuracy measure from the verbal response conflict task, but this could be explained by the fact that children's performance reached ceiling levels.

The probability of passing the DCCS also increased with age, yet even for children in the older age group (5–6 years) performance did not reach maximal accuracy (only 58% of 5-year olds achieved perfect post-switch performance. Interestingly, we found poorer levels of performance employing an emotional shifting task compared to previous results with the standard version of the DCCS in this age range. We believe that the explanation might relate to either (1) the greater impact of emotional expression as a categorizing criterion and in the resulting *negative priming effect,* or (2) the greater *perceptual conflict* between the two dimensions (color and emotional expression) induced by our stimuli. Related to the first explanation, Müller and Zelazo (2001) have proposed that a negative priming effect might be generated in the DCCS task by the need to inhibit a dimension (here, the emotional expression) in order to focus solely on the target dimension (i.e., color), and then to "undo" this initial inhibition during the second phase (i.e., when emotional expression becomes the target dimension). To be more specific, it is not that children have trouble with inhibiting this dominant dimension (in the pre-shift phase), but rather that they have difficulty disengaging this negative priming effect from the pre-shift set during the post-shift phase (Garon et al., 2008). This negative priming explanation could be tested in a future study by reversing the order of the dimensions (asking the child to categorize the items first by emotion, and then by color) which should theoretically reduce this effect. A second possible explanation relates to the higher degree of *perceptual conflict* elicited by the two dimensions during the pre-shift phase. The main distinction from the previous account is that it does not imply that sorting according to emotional expression was more salient, but that the target cards, were perceptually similar to a greater degree than, for instance, the boats and the rabbits. Apart from the color dimension which was clearly different, the emotional expression was related to a simple perceptual difference in the orientation of the mouth line. Future studies taking this explanation into account, could require the children to sort the cards according to the same two dimensions in the absence of the target cards, which has already been shown to improve performance (Perner and Lang, 2002), as no perceptual mismatch would be present. An alternative would be to separate the dimensions by placing them side by side on the card (as in Kloo and Perner, 2005).

CONCURRENT AND PREDICTIVE EFFECTS

The results suggest that given only a time difference of 9 months between measurements, the overlap in STM spans was sufficiently large, such that adding concurrent STM performance to a model already containing the previous STM did not improve EF prediction to a significant extent. On one hand, this indicates the stability of the predictive relationships between STM and WM measures. It is possible, however, that given a larger time difference between measurements, a direct effect could have been observed. However, for visuospatial WM, the addition of the second time point STM was significant, suggesting that the impact of this variable during this 9 months interval is not fully accounted by its previous development.

In the final models, STM at the first assessment was the most consistent predictor of performance across the EF measures. The best model in terms of predictive ability was the verbal WM, where the variables in the model accounted for over 50% of variance in the outcome, about a third of explanatory power being attributed to verbal STM. The models for visuospatial WM and shifting had a somewhat poorer predictive performance (only about 30–35% of variance was explained) and models for inhibition were inadequate for prediction purposes (20% or less of variance was explained). These are also the only models in which STM was a weak predictor, especially for response conflict, which diverged from previous findings by Espy and Bull (2005). However, it is important to note that in that study, children were divided into dichotomous High and Low digit span groups, while here a more refined continuous measure was used for both verbal and visuospatial STM performance. In our case, high verbal STM spans were indicative of good motor response conflict and shifting performance, while there were no links between STM spans and response suppression or verbal response conflict. While the associations between verbal STM span and motor response conflict parallel those obtained by Espy and Bull (2005), it is difficult to relate the results concerning the (absence of) associations between visuospatial STM, response suppression, and verbal response conflict to previous literature since the current experimental design is not directly comparable to any previous study with preschoolers. Hence, our results need to be validated in other samples before an explanation could be advanced. Also, the identified relationship between verbal STM and shifting performance warrants further exploration, suggesting that cognitive flexibility - as reflected by the Em-DCCS—might be strongly dependent on children's ability to verbally encode and maintain relevant stimulus-related information for brief successive periods of time. Preliminary evidence supporting this idea comes from the same study of Espy and Bull (2005), in which preschoolers with higher memory spans outperformed those with lower memory spans in the flexibility condition of the Shape School task. In a more systematic investigation of the contribution of WM (actually measured with a verbal span task) to the costs of cognitive flexibility in preschoolers, Chevalier and collaborators (2012) showed that after 4½ years of age, verbal STM was associated with specific costs on the same Shape School task. This evidence was related to the crucial role of verbal memory in the identification and maintenance of task goals necessary for performance on the flexibility tasks (Blaye and Chevalier, 2011).

THE ROLE OF INDIVIDUAL DIFFERENCES

Regarding *gender differences*, the only outcome on which such effects were found was shifting, as girls were significantly more likely to pass the DCCS task than boys. These results apparently are at odds with studies in which no gender differences were found in preschool children on the standard DCCS (e.g., Coldren and Colombo, 2009; Moriguchi et al., 2012). However, it is notable that some studies such as the one conducted by Wiebe

and collaborators (2008) did find evidence of higher absolute levels of EF performance in preschool girls. Our results are also in line with studies reporting that preschool girls presented higher levels of effortful control (e.g., Olson et al., 2005; Raaijmakers et al., 2008). Also, the fact that our task involved operating with emotional material (categorizing a stimulus based on emotion) might have favored the performance of preschool girls, as this has been previously indicated by their better performance at decoding emotion from facial expressions (Boyatzis et al., 1993) and their faster emotional judgments after neutral ones in a shifting context (Mocan et al., 2014).

Non-verbal intelligence scores were linked to superior WM spans, but did not have an impact on other EF performances. These results correspond to many adult and developmental studies showing that WM performance is closely related to intelligence scores (e.g., Fry and Hale, 2000; Colom et al., 2003), as both types of tasks employ attentional control (Engle, 2010). Moreover, non-verbal intelligence was modestly correlated with STM spans, but not with larger gains in children's STM spans. These results are in line with developmental findings suggesting that when the common variance between WM and STM is controlled, the residual WM factor is linked to children's intelligence, whereas the residual STM factor is not (Engel de Abreu et al., 2010). Hence, the high inter-subject variation observed for WM could be a reflection of the fact that WM performance relies on individual differences beyond STM performance. Taken together, our findings suggest that the link between intelligence and WM performance in young children could be mainly explained by the cognitive control mechanisms employed in WM tasks, and not by the storage component of such measures.

Regarding the impact of *trait anxiety* upon children's EFs, our hypotheses were partially confirmed. It is important to note that verbal STM performance at either time point correlated with initial levels of anxiety, and there was no significant link between anxiety and visuospatial STM. These results resonate well with the lack of anxiety-related effects on visuospatial STM found in a previous study with preschoolers (Visu-Petra et al., 2011). ACT (Eysenck et al., 2007) predicts that such effects should be less visible on the accuracy scores of tasks employing lower levels of attentional control, and more evident in efficiency measures, which were not available for our STM measurements. In line with this prediction and our current results, Visu-Petra and collaborators (Study 1, 2011) found no impact of trait anxiety upon preschoolers' STM accuracy, yet on the verbal storage tasks, there was a detrimental effect upon children's processing efficiency (i.e., duration of preparatory intervals). Moreover, it is important to note that higher anxiety at the start of the experiment did not result in lesser gains in STM performance, measured as the difference between the two STM measures.

With regards to children's performance on a task with similar levels of executive demands, we found that response suppression scores were not significantly affected by anxiety. However, contrary to our expectations, anxiety did not have a negative effect on either of the two response conflict measures, which presented higher levels of executive demands. There is very limited empirical evidence for such a relationship during early childhood for typically developing children (e.g., Cheie et al., 2014), as the studies have been mostly conducted with pediatric anxiety (e.g., Mueller et al., 2012). However, our negative findings should be regarded with caution considering the high levels of performance registered on all inhibitory control measures, as well as the fact that two of these measures provided only accuracy, and no efficiency outcomes.

As anticipated, trait anxiety was a significant predictor for WM components. The association between trait anxiety and verbal WM was significant even when controlling for other individual differences and STM spans. While the negative effect of anxiety was non-significant in the final model for visuospatial WM, the fact that anxiety significantly predicted performance in a previous model in which verbal (i.e., domain non-specific) STM was omitted is suggestive of the importance of this predictor for visuospatial WM. However, based on the magnitude of the effects, it is most likely that anxiety only affects WM to a practically relevant extent if the children are situated toward the upper end of the non-clinical spectrum. These results are in line with the ACT's predictions (Eysenck et al., 2007) regarding anxiety's deleterious impact upon updating, and correspond to the developmental empirical evidence highlighting such effects (e.g., Hadwin et al., 2005; Owens et al., 2008; Ng and Lee, 2010; Visu-Petra et al., 2011). To our knowledge, it is the first time that a detrimental impact of anxiety on a visuospatial WM measure is observed in young children. These findings are at odds with Visu-Petra and collaborators' (2011) results, which revealed no significant impact of anxiety upon young children's visuospatial updating performance. Yet, in adults, several studies have found individual differences in (threat-induced) state anxiety to account for performance variations in visuospatial WM (e.g., Shackman et al., 2006). Given the limited literature regarding the anxiety-visuospatial WM relationship in young children, replications of this effect are needed to shed light in this specific domain, especially considering the abovementioned idea of the higher executive load (and increased difficulty) experienced by children when performing the visuospatial memory tasks.

Anxiety also impacted preschoolers' shifting performance when controlling for individual differences in age, gender, and non-verbal intelligence. However, the effect became non-significant with the addition of STM measures. This result apparently fails to confirm our hypothesis and that of the ACT in predicting anxiety's detrimental effects in tasks employing set-shifting (Eysenck et al., 2007). On the other hand, taking a closer look at stimulus valence, the impact of anxiety on performance in the post-switching phase was restricted to children's perseverative errors in categorizing the sad (but not happy) faces according to the previous dimension (color). There is a documented general happy face advantage in recognizing even schematic facial expressions (Kirita and Endo, 1995), already visible in infants (Barrera and Maurer, 1981), which might have aided children's performance on this type of stimuli. However, we failed to replicate the facilitative effect of positive faces found in the study of Qu and Zelazo (2007). One crucial difference is that in the study by Qu and Zelazo (2007), in the happy/sad/neutral faces conditions, children were not required to perform any judgment based on emotion, but solely on age and gender. Therefore, the emotion

of the face was not the target of the evaluation, as it was in the current study. What could have impaired high-anxious children's performance in assessing the sad faces according to emotion, and made them continue in sorting them according to color? One important clue could come from the systematic analysis, performed by Kirita and Endo (1995) of how emotion displayed by schematic faces is recognized. Their study indicated that while happy (schematic) faces appeared to be recognized holistically, sad faces were more likely to be recognized in an analytical mode. In this respect, their results showed that sad faces were less disrupted by being presented in an inverted mode, as compared to the happy faces, for which the advantage completely disappeared in this inverted mode. It is plausible that this analytical mode of processing in recognizing emotion might have imposed greater executive demands, which have selectively disrupted high-anxious children's shifting performance. An alternative explanation would relate to their specific processing of negative emotional information conveyed by a sad face which would lead to a phenomenon of "cognitive avoidance" (Cloitre and Liebowitz, 1991) and probably to a re-focusing on perceptual aspects of performance such as stimulus color. However, a replication of this effect in an independent sample is required before attempting to distinguish between such potential explanations. Taken together, our findings reveal the crucial importance of taking individual differences (gender, intelligence, trait anxiety) into account when studying EF in young children, considering that such differences might influence, and might themselves be influenced, by individual progress in executive performance.

LIMITATIONS

There are several limitations which call into question the generalizability of our findings. Some of the limitations are methodological, induced by the study design, sample and procedure, while others are more related to the analytical approach—itself limited by the methodological constraints. More specifically, one of the main methodological limitations induced by looking at a developmental period characterized by intensive changes in all the assessed dimensions is that performance for the older children will inevitably reach *ceiling* levels of performance. This effect was found in our study to affect mostly measures of inhibition and shifting, similar to previous findings over the same age range (e.g., Lerner, 2012; Willoughby et al., 2012b). Another important methodological limitation was induced by the lack of a processing speed measure, this variable being causally related to changes in both memory span and executive functioning (Kail and Salthouse, 1994; Salthouse et al., 1998; Chuah and Mayberry, 1999).

A significant limitation makes us cautious with regards to directly incorporating current results in the ACT framework (Eysenck et al., 2007). As the theory predicts that anxiety-related worrisome thoughts interfere with the current task performance of an individual, the absence of a direct state anxiety measure (at both T1 and T2) precludes us from having clear-cut conclusions in this respect. However, using just a trait anxiety measure can be explained by the fact that in preschoolers, self-report measures of state anxiety are difficult to obtain in a reliable manner (Schniering et al., 2000). At the same time, studies also report that individuals with high levels of trait anxiety also experience higher levels of state anxiety in potentially stressful situations, such as performance evaluation for cognitive tasks (Lau et al., 2006). Nevertheless, trait anxiety was only evaluated at T1 and, while it could have remained stable within the T1-T2 interval, this was not directly verified in our study, this jeopardizing the incorporation of our findings in the ACT framework.

Considering the limitations of our analytic approach, it is important to note when discussing correlations among EF measures that these can arise from true similarities in the mechanisms underlying performance, but can also be confounded by common age-related effects and by shared method variance which can lead to spurious overlaps (e.g., reliance on verbal skills or on processing speed). These can only be eliminated by using multiple tasks tapping the same construct and relying on latent variable analysis to exclude such measurement error (Willoughby et al., 2012b). We only accomplished this objective to a certain extent, especially in terms of STM and WM measures, but to a lesser extent in terms of inhibition and especially of shifting performance. Also, other impediments eliminated the possibility of a latent variable analysis were the array of distributions of variables (ceiling effects noted above), and the limited sample size available at both time points.

CONCLUSIONS AND FUTURE DIRECTIONS

The aforementioned limitations notwithstanding, there are some particular strengths of the current study. These are reflected by the use of repeated assessments conducted at a young age and related to subsequent levels of performance, the choice (wherever possible) of multiple tasks to assess each construct and its subcomponents, and the often overlooked analyses of the impact of individual differences (in age, non-verbal intelligence, and anxiety).

First, regarding the developmental prerequisites of EF, STM appears to be a reliable and stable predictor during this interval (especially of WM and especially for the same stimulus modality). A cautionary note relates to several studies with preschoolers which have investigated WM by using tasks purported to measure STM (Hughes and Ensor, 2007; Wiebe et al., 2008; Noël, 2009). Very early during development such an overlap might be justified by the high demands posed by a memory span task for very young children (Reznick, 2007). However in older preschoolers our study concurs with other investigations (e.g., Alloway et al., 2006; Lerner, 2012) in revealing the necessity to delineate between STM tasks and WM tasks and to focus on the latter as a more adequate measure of EF. Repeated assessments of STM are necessary in order to identify the potential dynamics of this interrelationship and their presumed common reliance on attention control/processing speed improvements. We did not replicate the postulated distinction between response suppression and attention control mechanisms. It could be that at a young age they are truly undifferentiated, or that our tasks failed to impose a similar level of difficulty required in order to analyze inter-task correlations (see also Carlson, 2005).

Regarding the impact of individual differences, we found specific links between gender and shifting, between non-verbal intelligence and WM, and a potential link between trait anxiety

and verbal/visuospatial WM. While some of these results fit nicely and extend the theoretical frameworks proposed in the literature with adults (e.g., the ACT; Eysenck et al., 2007), they need substantial replication in larger independent samples and repeated assessments of individual differences over time. For instance, it would be relevant to measure anxiety at more than one time point in order to observe if past levels of anxiety affect performance beyond current levels, suggesting an early impact of anxiety on information-processing patterns, probably as a consequence of the enhanced plasticity of young children's threat-processing circuitry (Pine, 2007). Again, more intermediary time points are also needed in order to fully grasp the reciprocal interactions between cognitive, emotional, and (pre)dispositional factors during early development.

ACKNOWLEDGMENTS

This work was supported by two grants from the Romanian National Authority for Scientific Research, CNCS—UEFISCDI, project numbers PNII-RU-TE-2012-3-0323 and PN-II-ID-PCE-2012-4-0668. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors are grateful to Irina Bulai for her help with the data collection and to Paul Whitehead for proofreading the manuscript. Finally, we are thankful to the children, parents and kindergartens for their involvement in the study.

SUPPLEMENTARY MATERIALS

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00443/abstract

REFERENCES

Alloway, T. P. (2007). *Automated Working Memory Assessment*. London: Pearson Assessment.

Alloway, T. P., Gathercole, S. E., and Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: are they separable? *Child Dev.* 77, 1698–1716. doi: 10.1111/j.1467-8624.2006.00968.x

Anderson, P., Anderson, V., Northam, E., and Taylor, H. G. (2000). Standardization of the Contingency Naming Test (CNT) for school-aged children: a measure of reactive flexibility. *Clin. Neurophysiol. Assesment* 1, 247–273.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/S1364-6613(00)01538-2

Baddeley, A. D., and Hitch, G. J. (1974). "Working memory," in *The Psychology of Learning and Motivation*, ed G. H. Bower (New York, NY: Academic Press), 47–89.

Barrera, M., and Maurer, D. (1981). Recognition of mother's photographed face by the three-month-old infant. *Child Dev*. 52, 714–716. doi: 10.2307/1129196

Barkley, R. A. (1997). *ADHD and the Nature of Self-Control*. New York: Guilford Press.

Barrouillet, P., Bernardin, S., and Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *J. Exp. Psychol.* 133, 83–100. doi: 10.1037/0096-3445.133.1.83

Bayliss, D. M., Jarrold, C., Baddeley, A. D., and Gunn, D. M. (2005). The relationship between short–term memory and working memory: complex span made simple? *Memory* 13, 414–421. doi: 10.1080/09658210344000332

Bayliss, D. M., Jarrold, C., Gunn, D. M., and Baddeley, A. D. (2003). The complexities of complex span: explaining individual differences in working memory in children and adults. *J. Exp. Psychol.* 132, 71–92. doi: 10.1037/0096-3445.132.1.71

Benga, O., Ţincaş, I., and Visu-Petra, L. (2010). Investigating the structure of anxiety symptoms among Romanian preschoolers using the Spence Preschool Anxiety Scales. *Cogn. Brain Behav.* 14, 159–182.

Blaye, A., and Chevalier, N. (2011). The role of goal representation in preschoolers' flexibility and inhibition. *J. Exp. Child Psychol.* 108, 469–483. doi: 10.1016/j.jecp.2010.09.006

Boyatzis, C. J., Chazan, E., and Ting, C. Z. (1993). Preschool children's decoding of facial emotions. *J. Genet. Pers.* 154, 375–382. doi: 10.1080/00221325.1993.10532190

Brocki, K. C., Eninger, L., Thorell, L. B., and Bohlin, G. (2010). Interrelations between executive function and symptoms of hyperactivity/impulsivity and inattention in preschoolers: a two year longitudinal study. *J. Abnorm. Child. Psych.* 38, 163–171. doi: 10.1007/s10802-009-9354-9

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M., Moses, L. J., and Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Dev.* 69, 672–691. doi: 10.1111/j.1467-8624.1998.00672.x

Case, R. (1985). *Intellectual Development: Birth to Adulthood*. New York, NY: Academic Press.

Cheie, L., Veraksa, A., Zinchenko, Y., Gorovaya, A., and Visu-Petra, L. (2014). A cross-cultural investigation of inhibitory control, generative fluency, and anxiety symptoms in Romanian and Russian preschoolers. *Child Neuropsychol.* doi: 10.1080/09297049.2013.879111. [Epub ahead of print].

Chen, N. T., Clarke, P. J., MacLeod, C., and Guastella, A. J. (2012). Biased attentional processing of positive stimuli in social anxiety disorder: an eye movement study. *Cogn. Behav. Ther.* 41, 96–107. doi: 10.1080/16506073.2012.666562

Chevalier, N., Sheffield, T. D., Nelson, J. M., Clark, C. A., Wiebe, S. A., and Espy, K. A. (2012). Underpinnings of the costs of flexibility in preschool children: the roles of inhibition and working memory. *Dev. Neuropsychol.* 37, 99–118. doi: 10.1080/87565641.2011.632458

Chuah, Y. M. L., and Mayberry, M. T. (1999). Verbal and spatial short-term memory: common sources of developmental change? *J. Exp. Child Psychol.* 73, 7–44. doi: 10.1006/jecp.1999.2493

Ciairano, S., Visu-Petra, L., and Settanni, M. (2007). Executive inhibitory control and cooperative behavior during early school years: a follow-up study. *J. Abnorm. Child Psych.* 17, 335–345. doi: 10.1007/s10802-006-9094-z

Cloitre, M., and Liebowitz, M. R. (1991). Memory bias in panic disorder: an investigation of the cognitive avoidance hypothesis. *Cogn. Ther. Res.* 15, 371–386. doi: 10.1007/BF01173032

Coldren, J. T., and Colombo, J. (2009). Attention as a cueing function during kindergarten children's dimensional change task performance. *Infant Child Dev.* 18, 441–454. doi: 10.1002/icd.632

Colom, R., Flores-Mendoza, C., and Rebollo, I. (2003). Working memory and intelligence. *Pers. Indiv. Differ.* 34, 33–39. doi: 10.1016/S0191-8869(02)00023-5

Corsi, A. T. (1972). *Human Memory and the Medial Temporal Region of the Brain*. Montreal: McGill University.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185. doi: 10.1017/S0140525X01003922

Cowan, N., Nugent, L. D., Elliott, E. M., Ponomarev, I., and Saults, J. S. (1999). The role of attention in the development of short−term memory: age differences in the verbal span of apprehension. *Child Dev.* 70, 1082–1097. doi: 10.1111/1467-8624.00080

Cowan, N., Towse, J. N., Hamilton, Z., Saults, J. S., Elliott, E. M., Lacey, J. F., et al. (2003). Children's working-memory processes: a response-timing analysis. *J. Exp. Psychol.* 132, 113–132. doi: 10.1037/0096-3445.132.1.113

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: toward a unified theory of cognitive development and aging. *Dev. Rev.* 12, 45–75. doi: 10.1016/0273-2297(92)90003-K

Diamond, A. (1991). "Neuropsychological insights into the meaning of object concept development," in *Theepigenesis of Mind: Essays on Biology and Knowledge,* eds S. Carey and R. Gelman, (Hillsdale, NJ: Erlbaum), 67–110.

Diamond, A., and Gilbert, J. (1989). Development as progressive inhibitory control of action: retrieval of a contiguous object. *Cogn. Dev.* 4, 223–249. doi: 10.1016/0885-2014(89)90007-5

Diamond, A., and Kirkham, N. (2001). "Card sorting by children of 3 and 4 years and task switching by older children: inhibition needed to overcome attentional inertia," in *Rule Use Through the Lens of the Dimensional Change Card Sort: What Develops? Symposium Conducted at the Meeting of the Cognitive Development Society,* (Virginia Beach, VA).

Diamond, A., and Kirkham, N. (2005). Not as grown-up as we like to think: parallels between cognition in childhood and adulthood. *Psychol. Sci.* 16, 291–297. doi: 10.1111/j.0956-7976.2005.01530.x

Diamond, A., Carlson, S. M., and Beck, D. M. (2005). Preschool children's performance in task switching on the dimensional change card sort task: separating the dimensions aids the ability to switch. *Dev. Neuropsychol.* 28, 689–729. doi: 10.1207/s15326942dn2802_7

Engel de Abreu, P. M., Conway, A. R., and Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence* 38, 552–561. doi: 10.1016/j.intell.2010.07.003

Engle, R. W. (2010). Role of working memory capacity in cognitive control.*Curr. Anthropol.* 51, 17–26. doi: 10.1086/650572

Engle, R. W. (2002). Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1111/1467-8721.00160

Engle, R. W., Kane, M. J., and Tuholski, S. W. (1999). "Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex," in *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, eds A. Miyake and P. Shah (New York, NY: Cambridge University Press), 102–134.

Espy, K. A., and Bull, R. B. (2005). Inhibitory processes in young children and individual variation in short-term memory. *Dev. Neuropsychol.* 28, 669–688. doi: 10.1207/s15326942dn2802_6

Eysenck, M. W., and Derakshan, N. (2011). New perspectives in attentional control theory. *Pers. Indiv. Differ.* 50, 955–960. doi: 10.1016/j.paid.2010.08.019

Eysenck, M. W., Derakshan, N., Santos, R., and Calvo, M. (2007). Anxiety and cognitive performance: The Attentional Control Theory. *Emotion* 7, 336–353. doi: 10.1037/1528-3542.7.2.336

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent variable analysis. *J. Exp. Psychol.* 133, 101–135. doi: 10.1037/0096-3445.133.1.101

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *J. Exp. Psychol. Gen.* 137, 201–225. doi: 10.1037/0096-3445.137.2.201

Fry, A. F., and Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biol. Psychol.* 54, 1–34. doi: 10.1016/S0301-0511(00)00051-X

Garon, N., Bryson, S. E., and Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi: 10.1037/0033-2909.134.1.31

Gathercole, S. E., Willis, C. S., Emslie, H., and Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: a longitudinal study. *Dev. Psychol.* 28, 887–898. doi: 10.1037/0012-1649.28.5.887

Gathercole, S. E., Pickering, S. J., Ambridge, B., and Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* 40, 177–190. doi: 10.1037/0012-1649.40.2.177

Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action: performance of children 312–7 years old on a stroop-like day-night test. *Cognition* 53, 129–153. doi: 10.1016/0010-0277(94)90068-X

Hadwin, J. A., and Field, A. P. (2010). *Information Processing Biases and Anxiety: A Developmental Perspective.* Chichester: Wiley. doi: 10.1002/9780470661468

Hadwin, J. A., Donnelly, N., French, C. C., Richards, A., Watts, A., and Daley, D. (2003). The influence of children's self−report trait anxiety and depression on visual search for emotional faces. *J. Child. Psychol. Psyc.* 44, 432–444. doi: 10.1111/1469-7610.00133

Hadwin, J., Brogan, J., and Stevenson, J. (2005). State anxiety and working memory in children: a test of processing efficiency theory. *Educ. Psychol.* 25, 379–393. doi: 10.1080/01443410500041607

Hughes, C., and Ensor, R. (2007). Executive function and theory of mind: predictive relations from ages 2 to 4. *Dev. Psychol.* 43, 1447–1459. doi: 10.1037/0012-1649.43.6.1447

Hughes, C., and Ensor, R. (2011). Individual differences in growth in executive function across the transition to school predict externalizing and internalizing behaviors and self-perceived academic success at 6 years of age. *J. Exp. Child Psychol.* 108, 663–676. doi: 10.1016/j.jecp.2010.06.005

Hughes, C., Ensor, R., Wilson, A., and Graham, A. (2010). Tracking executive function across the transition to school: a latent variable approach. *Dev. Neuropsychol.* 35, 20–36. doi: 10.1080/87565640903325691

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Ingram, R. E., and Price, J. M. (2010). "Understanding psychopathology: the role of vulnerability," in *Vulnerability to Psychopathology. Risk across the Lifespan*, eds R. E. Ingram and J. M. Price (New York, NY: Guilford Publications), 3–17.

Jacques, S., and Zelazo, P. D. (2001). The flexible item selection task (FIST): A measure of executive function in preschoolers. *Dev. Neuropsychol.* 20, 573–591. doi: 10.1207/S15326942DN2003_2

Kail, R., and Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychol.* 86, 199–225. doi: 10.1016/0001-6918(94)90003-5

Kaplan, E., Fein, D., Morris, R., and Delis, D. (1991). *WAIS-R-NI Manual*. San Antonio, TX: Psychological Corporation.

Kennett, S., Eimer, M., Spence, C., and Driver, J. (2001). Tactile-visual links in exogenous spatial attention under different postures: convergent evidence from psychophysics and ERPs. *J. Cogn. Neurosci.* 13, 462–478. doi: 10.1162/08989290152001899

Kirita, T., and Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychol.* 89, 149–163. doi: 10.1016/0001-6918(94)00021-8

Klenberg, L., Korkman, M., and Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year old Finnish children. *Dev. Neuropsychol.* 20, 407–428. doi: 10.1207/S15326942DN2001_6

Kloo, D., and Perner, J. (2005). Disentangling dimensions in the dimensional change card−sorting task. *Dev. Sci.* 8, 44–56. doi: 10.1111/j.1467-7687.2005.00392.x

Korkman, M., Kirk, U., and Kemp, S. (1998). *NEPSY: A Developmental Neuropsychological Assessment.* New York, NY: The Psychological Corporation.

Lau, J. Y. F., Eley, T. C., and Stevenson, J. (2006). Examining the state–trait anxiety relationship: a behavioural genetic approach. *J. Abnorm. Child. Psych.* 34, 19–27. doi: 10.1007/s10802-005-9006-7

Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M. N., and Bull, R. (2012). The cognitive underpinnings of emerging mathematical skills: executive functioning, patterns, numeracy, and arithmetic. *Br. J. Educ. Psychol.* 82, 82–99. doi: 10.1111/j.2044-8279.2010.02016.x

Lehto, J. E., Juujärvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Lerner, M. D. (2012). *Executive Function Among Preschool Children: Unitary Versus Distinct Abilities.* Electronic Theses, Treatises and Dissertations. Paper 5389. Available online at: http://diginole.lib.fsu.edu/etd/5389.

Logie, R. H., and Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial working memory: evidence from developmental fractionation. *Eur. J. Cogn. Psychol.* 9, 241–257. doi: 10.1080/713752559

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203. doi: 10.1037/0033-2909.109.2.163

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., and Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *J. Cogn. Dev.* 13, 395–423. doi: 10.1080/15248372.2011.585478

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Pscyhol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., and Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? a latent-variable analysis. *J. Exp. Psychol. Gen.* 130, 621–640. doi: 10.1037/0096-3445.130.4.621

Mocan, O., Stanciu, O., and Visu-Petra, L. (2014). Attentional set-shifting in children: effects of individual differences in anxiety and attentional control. *Anxiety Stress Coping.* doi: 10.1080/10615806.2014.888419. [Epub ahead of print].

Moriguchi, Y., Evans, A. D., Hiraki, K., Itakura, S., and Lee, K. (2012). Cultural differences in the development of cognitive shifting: East–West comparison. *J. Exp. Child Psychol.* 111, 156–163. doi: 10.1016/j.jecp.2011.09.001

Morton, J. B., and Munakata, Y. (2002). Active versus latent representations: a model of perseveration, knowledge-action dissociation, and décalage. *Dev. Psychobiol.* 40, 255–265. doi: 10.1002/dev.10033

Mueller, S. C. (2011). The influence of emotion on cognitive control: relevance for development and adolescent psychopathology. *Front. Psychol.* 2:327, doi: 10.3389/fpsyg.2011.00327

Mueller, S. C., Hardin, M. G., Mogg, K., Benson, V., Bradley, B. P., Reinholdt-Dunne, M. L., et al. (2012). The influence of emotional stimuli on attention orienting and inhibitory control in pediatric anxiety. *J. Child. Psychol. Psyc.* 53, 856–863. doi: 10.1111/j.1469-7610.2012.02541.x

Müller, U., and Zelazo, P.D. (2001). "The role of selective attention and negative priming in the DCCS," in *Paper Presented at the Second Biennial Meeting of the Cognitive Development Society,* (Virginia Beach).

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends Cogn. Sci.* 5, 309–315. doi: 10.1016/S1364-6613(00)01682-X

Ng, E. L., and Lee, K. (2010). Children's task performance under stress and non-stress conditions. *Cogn. Emot.* 24, 1229–1238. doi: 10.1080/02699930903172328

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychol. Bull.* 126, 220–246. doi: 10.1037/0033-2909.126.2.220

Noël, M. P. (2009). Counting on working memory when learning to count and to add: a preschool study. *Dev. Psychol.* 45, 1630–1643. doi: 10.1037/a0016224

Olson, S. L., Sameroff, A. J., Kerr, D. C., Lopez, N. L., and Wellman, H. M. (2005). Developmental foundations of externalizing problems in young children: the role of effortful control. *Dev. Psychopatol.* 17, 25–45. doi: 10.1017/S0954579405050029

Owens, M., Stevenson, J., Hadwin, J. A., and Norgate, R. (2012). Anxiety and depression in academic performance: an exploration of the mediating factors of worry and working memory. *School Psychol. Int.* 33, 433–449. doi: 10.1177/0143034311427433

Owens, M., Stevenson, J., Norgate, R., and Hadwin, J. A. (2008). Processing efficiency theory in children: working memory as a mediator between trait anxiety and academic performance. *Anxiety Stress Coping* 21, 417–430. doi: 10.1080/10615800701847823

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychol.* 32, 301–345. doi: 10.1016/0001-6918(70)90108-3

Perner, J., and Lang, B. (2002). What causes 3−year−olds' difficulty on the dimensional change card sorting task?. *Infant Child Dev.* 11, 93–105. doi: 10.1002/icd.299

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* 9, 148–158. doi: 10.1038/nrn2317

Pickering, S. J., Gathercole, S. E., Hall, M., and Lloyd, S. A. (2001). Development of memory for pattern and path: further evidence for the fractionation of visuo-spatial working memory. *Q. J. Exp. Psychol.* 54A, 397–420. doi: 10.1080/713755973

Pickering, S. J., Gathercole, S. E., and Peaker, S. H. (1998). Verbal and visuo-spatial short-term memory in children: evidence for common and distinct mechanisms. *Mem. Cogn.* 26, 1117–1130. doi: 10.3758/BF03201189

Pine, D. S. (2007). Research review: a neuroscience framework for pediatric anxiety disorders. *J. Child. Psychol. Psyc.* 48, 631–648. doi: 10.1111/j.1469-7610.2007.01751.x

Posner, M. I., Rothbart, M. K., and Harman, C. (1994). *Emotion and Culture: Empirical Studies of Mutual Influence*. Washington, DC: American Psychological Association.

Qu, L., and Zelazo, P. D. (2007). The facilitative effect of positive stimuli on 3-year-old's flexible rule use. *Cogn. Dev.* 22, 456–473. doi: 10.1016/j.cogdev.2007.08.010

Raaijmakers, M. A., Smidts, D. P., Sergeant, J. A., Maassen, G. H., Posthumus, J. A., Van Engeland, H., et al. (2008). Executive functions in preschool children with aggressive behavior: impairments in inhibitory control. *J. Abnorm. Child Psych.* 36, 1097–1107. doi: 10.1007/s10802-008-9235-7

Raven, J., Raven, J. C., and Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 2: The Coloured Progressive Matrices*. Oxford: Oxford Psychologists Press.

Razza, R. A., and Blair, C. (2009). Associations among false-belief understanding, executive function, and social competence: a longitudinal analysis. *J. Appl. Dev. Psychol.* 30, 332–343. doi: 10.1016/j.appdev.2008.12.020

Reznick, J. S. (2007). "Working memory in infants and toddlers," in *Short- and Long-Term Memory in Infancy and Early Childhood: Taking the First Steps Toward Remembering*, eds L. M. Oakes and P. J. Bauer (Oxford: Oxford University Press), 3–26.

Riggs, N. R., Blair, C. B., and Greenberg, M. T. (2004). Concurrent and 2-year longitudinal relations between executive function and the behavior of 1st and 2nd grade children. *Child Neuropsychol.* 9, 267–276. doi: 10.1076/chin.9.4.267.23513

Roberts, R. J., and Pennington, B. F. (1996). An integrative framework for examining prefrontal cognitive processes. *Dev. Neuropsychol.* 12, 105–126. doi: 10.1080/87565649609540642

Röthlisberger, M., Neuenschwander, R., Cimeli, P., Michel, E., and Roebers, C. M. (2012). Improving executive functions in 5−and 6−year−olds: Evaluation of a small group intervention in prekindergarten and kindergarten children. *Infant Child Dev.* 21, 411–429. doi: 10.1002/icd.752

Russell, J., Jarrold, C., and Henry, L. (1996). Working memory in children with autism and with moderate learning difficulties. *J. Child. Psychol. Psyc.* 37, 673–686. doi: 10.1111/j.1469-7610.1996.tb01459.x

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology* 19, 532–545. doi: 10.1037/0894-4105.19.4.532

Salthouse, T. A., Fristoe, N., McGuthry, K. E., and Hambrick, D. Z. (1998). Relation of task switching to speed, age, and fluid intelligence. *Psychol. Aging.* 13, 445–461. doi: 10.1037/0882-7974.13.3.445

Schneider, W., Kron, V., Hünnerkopf, M., and Krajewski, K. (2004). The development of young children's memory strategies: first findings from the Würzburg Longitudinal Memory Study. *J. Exp. Child Psychol.* 88, 193–209. doi: 10.1016/j.jecp.2004.02.004

Schniering, C., Hudson, J. L., and Rapee, R. M. (2000). Issues in the assessment and diagnosis of anxiety disorders in children and adolescents. *Clin. Psychol. Rev.* 20, 453–478. doi: 10.1016/S0272-7358(99)00037-9

Spence, S. H., Rapee, R., McDonald, C., and Ingram, M. (2001). The structure of anxiety symptoms among preschoolers. *Behav. Res. Ther.* 39, 1293–1316. doi: 10.1016/S0005-7967(00)00098-X

Shackman, A. J., Sarinopoulos, I., Maxwell, J. S., Pizzagalli, D. A., Lavric A., and Davidson, R. J. (2006). Anxiety selectively disrupts visuospatial working memory. *Emotion* 6, 40–61. doi: 10.1037/1528-3542.6.1.40

Simpson, A., and Riggs, K. J. (2005). Inhibitory and working memory demands of the day–night task in children. *Br. J. Dev. Psychol.* 23, 471–486. doi: 10.1348/026151005X28712

Smith, E. E., and Jonides, J. (1997). Working memory: a view from neuroimaging. *Cogn. Pscyhol.* 33, 5–42. doi: 10.1006/cogp.1997.0658y

Thorell, L. B., and Wåhlstedt, C. (2006). Executive functioning deficits in relation to symptoms of ADHD and/or ODD in preschool children. *Infant Child Dev.* 15, 503–518. doi: 10.1002/icd.475

Tillman, C., Brocki, K. C., Sørensen, L., and Lundervold, A. J. (2013). A longitudinal examination of the developmental executive function hierarchy in children with externalizing behavior problems. *J. Atten. Disord*. doi: 10.1177/1087054713488439. [Epub ahead of print].

Van der Ven, S. H., Kroesbergen, E. H., Boom, J., and Leseman, P. P. (2012). The development of executive functions and early mathematics: a dynamic relationship. *Br. J. Educ. Psychol.* 82, 100–119. doi: 10.1111/j.2044-8279.2011.02035.x

Visu-Petra, L., Cheie, L., Benga, O., and Alloway, T. P. (2011). Effects of anxiety on memory storage and updating in young children. *Int. J. Behav. Dev.* 35, 38–47. doi: 10.1177/0165025410368945

Visu-Petra, L., Cheie, L., and Miu, A. C. (2013a). "Working memory and anxiety: exploring the interplay of individual differences across development," in *Working Memory: the Connected Intelligence*, eds T. P. Alloway and R. G. Alloway (New York, NY: Psychology Press), 187–216.

Visu-Petra, L., Cheie, L., and Mocan, O. (2013b). "Executive functioning in high-trait anxious children: a cognitive vulnerability factor?" in *Stress and Anxiety,* eds K. Moore, K. Kaniasty, and P. Buchwald (Berlin: Logos), 153–160.

Visu-Petra, L., Miclea, M., Cheie, L., and Benga, O. (2009). Processing efficiency in preschoolers' memory span: individual differences related to age and anxiety. *J. Exp. Child Psychol.* 103, 30–48. doi: 10.1016/j.jecp.2008.09.002

Visu-Petra, L., Ţincaş, I., Cheie, L., and Benga, O. (2010). Anxiety and visual-spatial memory updating in young children: an investigation using emotional facial expressions. *Cogn. Emot*. 24, 223–240. doi: 10.1080/02699930903387546

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children-(WISC-III)*. San Antonio, TX: Psychological Corporation.

Wiebe, S. A., Espy, K. A., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., and Espy, K. A. (2011). The structure of executive function in 3-year-olds. *J. Exp. Child Psychol.* 108, 436–452. doi: 10.1016/j.jecp.2010.08.008

Willoughby, M. T., Blair, C. B., Wirth, R. J., and Greenberg, M. (2010). The measurement of executive function at age 3 years: psychometric properties and criterion validity of a new battery of tasks. *Psychol. Assessment* 22, 306–317. doi: 10.1037/a0018708

Willoughby, M. T., Kupersmidt, J. B., and Voegler-Lee, M. E. (2012a). Is preschool executive function causally related to academic achievement? *Child Neuropsychol.* 18, 79–91. doi: 10.1080/09297049.2011.578572

Willoughby, M. T., Blair, C. B., Wirth, R. J., and Greenberg, M. (2012b). The measurement of executive function at age 5: psychometric properties and relationship to academic achievement. *Psychol. Assessment* 24, 226–239. doi: 10.1037/a0025361

Wright, I., Waterman, M., Prescott, H., and Murdoch−Eaton, D. (2003). A new Stroop−like measure of inhibitory function development: typical developmental trends. *J. Child. Psychol. Psyc.* 44, 561–575. doi: 10.1111/1469-7610.00145

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Zelazo, P. D., and Müller, U. (2002). The balance beam in the balance: reflections on rules, relational complexity, and developmental processes. *J. Exp. Child Psychol.* 81, 458–465. doi: 10.1006/jecp.2002.2667

Zelazo, P. D., Müller, U., Frye, D., and Marcovitch, S. (2003). The development of executive function in early childhood. *Monogr. Soc. Res. Child.* 68, 1–137. doi: 10.1111/j.0037-976X.2003.00261.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 January 2014; accepted: 26 April 2014; published online: 16 May 2014.
Citation: Visu-Petra L, Stanciu O, Benga O, Miclea M and Cheie L (2014) Longitudinal and concurrent links between memory span, anxiety symptoms, and subsequent executive functioning in young children. Front. Psychol. ***5****:443. doi: 10.3389/fpsyg.2014.00443*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Visu-Petra, Stanciu, Benga, Miclea and Cheie. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 30 April 2014
doi: 10.3389/fpsyg.2014.00381

Inhibitory processes in toddlers: a latent-variable approach

Elena Gandolfi*, Paola Viterbori, Laura Traverso and M. Carmen Usai

Department of Educational Science, University of Genoa, Genoa, Italy

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Kerry Lee, National Institute of Education, Singapore
Caron Ann Campbell Clark, University of Oregon, USA
Yoshifumi Ikeda, Tokyo Gakugei University, Japan

****Correspondence:***
Elena Gandolfi, Department of Educational Science, University of Genoa, Corso A. Podestà 2, 16128 Genoa, Italy
e-mail: elena.gandolfi@unige.it

The aim of this study was to investigate the nature of inhibitory processes in early childhood. A confirmatory factor analysis was used to examine the latent structure of inhibitory processes in day-care center children aged 24–32 months and in preschool children aged 36–48 months. The best fit to the data for the younger sample was a single undifferentiated inhibition factor model; in older children, a two-factor model was differently identified in which response inhibition and interference suppression were distinguished.

Keywords: inhibitory processes, executive functions, latent structure, early childhood, confirmatory factor analysis

INTRODUCTION

Inhibitory processes are considered important components of cognition that affect an individual's ability to function in everyday life. Using a broad definition, inhibitory processes refer to the ability to control one's mental processes and responses, to ignore an internal or an external prompt, and to perform an alternative action (Diamond, 2013).

During childhood, inhibitory processes have been found to affect different aspects of child functioning, such as self-regulatory behaviors (e.g., Rueda et al., 2005), theory of mind (e.g., Carlson et al., 2002), and the internalization of conduct standards (e.g., Kochanska et al., 1996). Inefficient inhibitory processes have been linked to several developmental disorders, such as attention deficit hyperactivity disorder (Barkley, 1997; Ozonoff and Jensen, 1999; Schachar et al., 2000), obsessive compulsive disorder, and autistic spectrum disorder (Ozonoff et al., 1991; Robinson et al., 2009).

In the recent literature, inhibitory control has been considered one aspect of the multi-component construct of executive function that proves to be clearly separate from other executive dimensions, such as shifting and working memory (WM), in adults (Miyake et al., 2000) and older children (Lehto et al., 2003). However, in younger children, the separability of different executive functions remains a matter of debate.

Using a confirmatory statistical approach, previous studies have found that a single undifferentiated executive control factor was the most appropriate means of describing the executive latent structure in early childhood and in preschoolers (Wiebe et al., 2008, 2011; Hughes et al., 2010; Willoughby et al., 2010; Fuhs and Day, 2011). Diverging from previous results, Miller et al. (2012) reported that a two-factor model, which consisted of WM and inhibition, fitted the data better in a sample of preschoolers between the ages of 3 and 5 years than did a single-factor model or a three-factor model composed of WM, inhibition, and shifting. Similarly, Usai et al. (2014) found that a two-factor model provided the best fit for the data, with inhibition as a separate dimension from a working memory-flexibility factor, at both 5 and 6 years of age. These studies suggest that an emerging differentiation of EF processes is already apparent in early childhood and that inhibitory processes emerge as a separate dimension as early as the preschool years.

Although a fairly large body of literature exists on inhibitory processes and their role in child functioning and development, a precise definition of inhibition remains elusive.

INHIBITION: SINGLE OR MULTIPLE PROCESSES?

An important shift in the research on inhibition concerns the idea that inhibition may be better conceptualized as a set of functions than as a unitary construct (Dempster, 1992; Nigg, 2000). Of course, this approach implies that there are commonalities as well as differences between the various inhibitory functions.

In his review, Nigg (2000) distinguished between the effortful inhibition of a motor or cognitive response and the automatic inhibition of attention. He included in the first category four different types of inhibition: (a) interference control, which is the ability to prevent interference due to resource or stimulus competition; (b) cognitive inhibition, which involves suppressing non-pertinent thoughts to preserve other processes, such as WM or attention; (c) behavioral inhibition, which refers to the ability to overcome a prepotent response or a prompted but socially inappropriate response; and (d) oculomotor inhibition, which involves suppressing a reflexive saccade.

The taxonomy of Nigg (2000) can be considered a theoretical attempt to describe the different inhibitory functions. Friedman and Miyake (2004), using a latent variable analysis, subsequently distinguished between three main forms of inhibition (prepotent response inhibition, resistance to distractor interference, and resistance to proactive interference), more or less in accordance with the taxonomy of Nigg (2000).

Prepotent response inhibition was defined as the ability to intentionally prevent a dominant, automatic, or prepotent response; resistance to distractor interference was identified as the

ability to overcome an interference that is external to the individual and irrelevant to the current task; and resistance to proactive interference was defined as the ability to control interference from previous tasks. In adults, the results suggested that the term "inhibition" could not be overextended to different processes. A common inhibition ability was found, which was represented by inhibition of action (prepotent response inhibition) and inhibition of attention (resistance to distractor interference), both of which involve the ability to actively maintain critical goal-related information. Surprisingly, resistance to proactive interference was unrelated to both the prepotent response inhibition and resistance to distractor interference, which suggests that this type of cognitive inhibition acts as an independent dimension.

More recently, in a review of the literature, Diamond (2013) suggested that inhibitory control could be divided into three main components: inhibition at the level of thought and memories (cognitive inhibition), inhibition at the level of attention (executive attention), and inhibition at the level of behavior (response inhibition). Cognitive inhibition and executive attention are the mechanisms underlying interference control, which involves both the ability to suppress interfering (or prepotent) mental representations and the ability to ignore (or inhibit attention to) particular stimuli to attend to other stimuli based on one's goals or intentions.

In contrast, the response inhibition component involves the ability to regulate one's behavior and control one's emotions to support the regulation of a behavior. This ability involves preventing impulsive behaviors when completing a task despite being faced with distractions or other competing stimuli. In children, this behavioral self-control is facilitated if there is sufficient time between the triggering stimulus and the response the child should produce (Simpson et al., 2012).

A distinction between the capacity to suppress prepotent but inappropriate responses (response inhibition) and the ability to filter out irrelevant information in the environment (interference monitoring and suppression) was also suggested by Bunge et al. (2002), who found differences in the regions of neural activation associated with response inhibition and interference suppression. This distinction is based on the differences between tasks that constitute potentially conflicting dimensions, such as the *Flanker* tasks, and univalent tasks, in which only a single feature is presented and the conflict is between two response options to the same stimulus. This situation creates a conflict between the habitual response and a less familiar arbitrary response, as in the *Day/Night Stroop* task. It has been suggested (Blasi et al., 2006) that in tasks requiring interference suppression, both the response conflict and the process of filtering out incongruent information within the stimulus are present.

INHIBITORY CONTROL DEVELOPMENT

The presence of inhibitory processes in toddlerhood and preschool-aged children has been established in many studies (Kochanska et al., 1997; Diamond, 2002; Jones et al., 2003; Carlson, 2005; Garon et al., 2008). However, given that the term inhibition has been variously used in the literature, as noted previously, it is not easy to extract a general trajectory of the development of inhibitory processes from the literature.

In their review, Best and Miller (2010) suggested that significant development of inhibitory processes occurs in the preschool years. By age 4, children show signs of successful performance on both response inhibition tasks and complex inhibition tasks, which require substantial WM. Inhibition continues to improve, especially from 5 to 8 years of age and particularly for tasks that combine inhibition and WM (Gerstadt et al., 1994; Carlson, 2005). According to Best and Miller (2010), these later improvements are unlikely to reflect fundamental cognitive changes, such as a preschooler's acquisition of the rule-formation ability, which is necessary for performing tasks such as the *Dimensional Change Card Sort*. Instead, the fundamental changes in cognition consist of quantitative improvements in accuracy.

Rueda et al. (2005), distinguishing response inhibition and inhibition in the attentional domain from conflict resolution, claimed that the ability to resolve conflict is the most important milestone in EF development, which develops slowly in the first 2 years of life and improves noticeably between 2 and 5 years of age. Similarly, Clark et al. (2013), using growth curve modeling to describe the growth trajectories of inhibitory control and cognitive flexibility, found a sizeable increase in these abilities, particularly during the period of 3–4 years of age. The authors suggested that this accelerated growth may reflect a qualitative change in executive processing. They also found differences in the developmental trajectories of different task conditions related to different cognitive demands.

Garon et al. (2008) distinguished between simple and complex inhibition processes, referring particularly to tasks that are employed to explore the inhibitory processes during early childhood, and classified them according to WM demands. These authors ascribed paradigms, such as the *Don't*, the *Delay gratification*, the *Object retrieval,* and the *Antisaccade*, to simple inhibition tasks. Conversely, they included the *Simon-like* tasks, the *Flanker* tasks, the *Less is more* task, the *Hand game* task, and the *Knock and tap* task as examples of complex inhibition paradigms because these tasks require the resolution of conflict between dominant and subdominant responses and, consequently, involve greater levels of top-down control. Best and Miller (2010) also included the *Dimensional Change Card Sort* in these complex tasks because it determines a prepotent response during the pre-switch phase that must later be inhibited. In the post-switch phase, the child is asked to sort the same cards by the other dimension that conflicts with the previous one, which remains visible.

THE PRESENT STUDY

The aim of the present study was to investigate the latent organization of inhibitory processes in early childhood. Following the hypothesis of Bunge et al. (2002), we considered two dimensions of inhibition: *response inhibition* with low WM demands and *interference suppression*, which is associated with higher WM demands and requires the individual to address interference or conflict from recently presented information or to filter out incongruent information. We performed a cross-sectional study in which a confirmatory factor analysis (CFA) was used to investigate the latent structure of inhibitory processes in children aged 24–32 and 36–48 months. Research investigating the underlying

construct of inhibition at these age levels is currently absent from the literature.

Two different models of inhibition were tested. First, we considered a unitary factor model based on earlier studies indicating that a single, undifferentiated, executive control factor was the most appropriate for describing the executive latent structure in preschoolers (Wiebe et al., 2008, 2011; Hughes et al., 2010). We subsequently examined a two-factor model in which *response inhibition* was distinguished from *interference suppression*. To test these two models, we chose measures that assessed the ability to suppress prepotent but inappropriate responses (*response inhibition*) and the ability to manage the interference of potentially conflicting features of the task (*interference suppression*), as suggested by Bunge et al. (2002).

In *response inhibition* tasks, the conflict is between two response options to the same stimulus, namely, the habitual response and a less familiar response. For example, the *Circle Drawing Task* and the *Tower Building* task require the ability to suppress an impulsive motor response when a task calls for it; similarly, in the *Bear/Dragon* task, the child needs to selectively suppress commanded actions in response to a stimulus based on a rule. In the *Day/Night Stroop* task, the child must suppress the tendency to produce a dominant response (say "day" when a card with a sun is presented) in favor of a subdominant response (say "night" when a card with a sun is presented). These tasks are examples of univalent displays in which only a single feature is presented and the conflict is between two response options to the same stimulus feature (Martin-Rhee and Bialystok, 2008).

Interference suppression tasks require the child to select a piece of information from a complex stimulus that is misleading and in which interfering features of the stimulus must be inhibited. These latter tasks involve greater levels of cognitive control, are associated with higher WM demands, and require individuals to filter out irrelevant information.

For example, in the *Fish task*, the child must respond to a central target flanked by distractors whose interference must be inhibited. In the *Reverse Categorization* task and the *Dimensional Change Card Sort* task, children must classify objects or cards by considering their different features, inhibiting the sorting rule previously learned. In particular, children must inhibit their attention to a dimension of the stimulus that was previously useful to solve the task and attend to a different aspect of the same stimulus. The *Animal House* task requires the child to match animal stickers with a color following a precise association rule and inhibiting the previous animal-color association each time.

To our knowledge, there has been no systematic attempt to empirically evaluate these dimensions of inhibition in early childhood.

MATERIALS AND METHODS

PARTICIPANTS

The present study involved two samples of 130 typically developing children: 60 children between the ages of 24 and 32 months (mean age = 28.41 months; $SD = 2.68$; $n = 25$ males and 35 females) in their last year of day-care and 70 children between the ages of 36 and 48 months (mean age = 42.35 months; $SD = 3.18$; $n = 34$ males and 36 females) in their first year of preschool. The participants were recruited by contacting six day-care centers and four kindergartens in the largest town in a northern region of Italy. Written parental informed consent was obtained before the participating children were admitted to the assessment sessions. Parents also completed a socioeconomic and educational background questionnaire: the mother's education level ranged from 8 to 18 years (mean = 13.7 years), and the father's education level ranged from 5 to 18 years (mean = 11.57 years); the mother's annual income ranged from 0 to 42,000 € (mean = 17,000 €), and the father's annual income ranged from 14,000 € to 42,000 € (mean = 22,000 €). Children with documented health problems, such as neurological, psychiatric or developmental disorders, or whose primary language spoken at home was not Italian were excluded from the study.

PROCEDURE

The children were tested individually in a quiet room of their day-care center or preschool during a 30- to 40-min session. Researchers and trained graduate students administered and scored all tests. A battery of inhibitory tasks, varying in format and response demands, were administered to the children in a standard order.

INHIBITORY MEASURES

A battery of tasks was employed to assess two inhibitory abilities: *response inhibition*, which is the ability to suppress a prepotent but inappropriate response to a stimulus, and *interference suppression*, which is the ability to address the interference of potentially conflicting characteristics of a stimulus.

The following measures were administered to children aged 24–32 months:

(1) Response inhibition:
The *Circle Drawing Task* (Bachorowski and Newman, 1985) assesses the ability to control an ongoing motor response. A circle is drawn on a cardboard square. The circle has a small arrow printed above its line to indicate the starting point and the direction of the tracing. The task is administered under two conditions: first with neutral instruction ("*Trace the circle with your finger*"), followed by an inhibition instruction ("*Trace the circle again, but this time as slowly as you can*"). The score is calculated as the proportion of the slowdown to the total time using the following formula: T1−T2/T1+T2, where T1 and T2 are the times recoded for the first and the second trials, respectively.
The *Tower Building* (Kochanska et al., 1996) evaluates the ability to take turns and to inhibit a prepotent response as in a go-no go task. The children are asked to take turns with the experimenter to build a tower using 20 wooden blocks (10 red and 10 blue). The score indicates the number of correct turns (range: 0–10).

(2) Interference suppression:
The *Fish Task* (Viterbori et al., 2012; adapted from Rueda et al., 2004) evaluates visual interference using an adaptation of the flanker paradigm (Eriksen and Eriksen, 1974). This is a forced-choice task in which children are required to point at where a centrally located target fish is oriented, despite the presence of interfering stimuli (other fishes) whose interference must be inhibited. There are 14 trials: 2 training trials,

6 congruent trials with the target and the interfering stimuli oriented in the same direction and 6 incongruent trials with the target and the interfering stimuli oriented in the opposite direction. Congruent and incongruent trials are randomly presented. The accuracy in the incongruent trials is scored (range: 0–6).

The *Animal House* (adapted from WPPSI; Wechsler, 1973) measures a child's ability to choose the correct association between stimuli (i.e., animal-color) by filtering out the other competing possibilities. The examiner shows the child three stickers, which represent a duck, a mouse and a frog, that are each matched with a colored house: the duck is matched with a red house, the mouse is matched with a blue house and the frog is matched with a yellow house. The child is then asked to correctly match 20 animal stickers (duck, mouse, and frog) with 20 different colored houses (red, blue, and yellow). In order to reduce WM load, before starting, the experimenter provided the child with an example of the matching rules which remained visible during the whole task. The score is obtained by calculating the total number of correctly matched stickers (range: 0–20).

The *Reverse Categorization* (Carlson et al., 2004) evaluates the ability to classify an object according to different rules. The task requires an individual to resolve a conflict generated by the previous presentation of a classification rule, which subsequently represents a source of interference. Children are introduced to two buckets and 12 blocks (six small and six big). The experimenter, using demonstration and verbal explanation, asks the child to sort big blocks into the "big" bucket and little blocks into the "little" bucket (pre-switch phase). Then, the experimenter reverses this categorization scheme (post switch phase) and suggests playing a "silly game" in which the children have to sort big blocks into the "small" bucket and small blocks into the "big" bucket. For each trial, the experimenter repeats the rule and then identifies the current block as big or small. There are 12 test trials for each phase, and no feedback is given. The score is the number of correct classifications in the post-switch phase (range: 0–12).

The following measures were administered to children aged 36–48 months:

(1) Response inhibition:

The *Circle Drawing Task* (Bachorowski and Newman, 1985) is the same as described above.

The *Bear/Dragon* (Reed et al., 1984) assesses the ability to inhibit or activate a motor response following a rule, in a similar way as in a *go no-go* task. The experimenter introduces children to a "nice" bear puppet and a "naughty" dragon puppet. The children are told that in this game, they are to do what the bear asks them to do (e.g., "touch your nose"), but not to do what the dragon asks. After practicing, there are 10 test trials with the bear and dragon commands in alternating order. The children are seated at a table throughout the task, and all actions involve hand movements. The performances on the bear and dragon trials are considered to be an index of self-control. The tasks are scored as follows: "0 indicates a movement or response when the dragon asks and no movement when the bear asks; 3 indicates no movement when the dragon asks and a movement or response when the bear asks" Also partial credits were scored: 2 indicates a partial movement or response when the bear asks, and a wrong movement when the dragon asks; 1 indicates a wrong movement when the bear asks and a partial movement or response when the dragon asks. The score ranges from 0 to 30.

The *Day/Night Stroop* (Gerstadt et al., 1994) assesses the ability to inhibit a prepotent verbal response and to activate an alternative verbal response. The experimenter presents a white card with a yellow sun and a black card with a white moon and stars on it. The children are instructed that in this game, they must say "Night" for the sun cards and "Day" for the moon cards. There are 16 test trials with each card presented in a fixed and pseudorandom order. There are no breaks or rule reminders. The accuracy (the number of correct items out of 16) is recorded (range: 0–16).

(2) Interference suppression:

The *Dimensional Change Card Sort* (DCCS, Zelazo, 2006) evaluates the extent to which young children between three and six years of age are able to remember two sets of rules, apply them and then switch the rules. This task requires children to address the interference generated by the previous sorting rule. Children are introduced to two recipe boxes, which have rectangular slots cut in the top. Target cards (a red rabbit and a blue boat) are affixed to the front of the boxes. The experimenter presents a series of cards (red and blue rabbits and boats) and instructs the children to place all the rabbits in the box with the red rabbit and all the boats in the box with the blue boat in the "shape game." After five consecutively correct trials, the experimenter asks the children to stop playing the "shape game" and to play the "color game" (post-switch phase). In this case, all the red items must go in the box with the red rabbit affixed, and all the blue items must go in the box with the blue boat affixed. In the third sorting phase (border phase), the experimenter explains that if there is a black border on a card, then the children must sort according to color; however, if there is no border, then they must sort according to shape. There are 24 trials (6 for the pre-switch phase, 6 for the post-switch phase, and 12 for the border phase); the score represents the number of correct responses (0–24).

The Fish Task (Viterbori et al., 2012; adapted from Rueda et al., 2004) and the Animal House task (adapted from WPPSI; Wechsler, 1973) are the same as described above.

RESULTS

DESCRIPTIVE STATISTICS

The descriptive statistics for all inhibitory measures are shown in **Table 1**.

No outliers were identified. The missing values for all measures ranged from 0 to 3%.

All dependent variables displayed adequate distributional characteristics, without substantial skewness or kurtosis. For both age range measures, skewness and kurtosis coefficients were

Table 1 | Descriptive statistics for the inhibitory measures used in 24–32 month sample and in 36–48 month sample.

	24–32 month sample						36–48 month sample				
	N	**Mean**	***SD***	**Skewness**	**Kurtosis**		***N***	**Mean**	***SD***	**Skewness**	**Kurtosis**
RESPONSE INHIBITION											
Circle	56	0.18	0.23	−0.07	0.51	Circle	70	0.33	0.24	−0.09	−0.56
Tower	56	5.35	2.8	−0.11	−0.73	Bear	69	19.79	6.62	0.41	−1.05
–	–	–	–	–	–	D/N	70	11.28	4.8	−1.21	0.49
INTERFERENCE SUPPRESSION											
Fish	57	2.61	1.67	−0.05	−0.79	Fish	70	4.45	1.85	−0.91	−0.49
Animal	60	9.47	3.75	0.77	1.12	Animal	70	18.00	3.33	1.98	3.13
Reverse	60	9.65	2.74	−0.75	0.9	DCCS	70	17.27	4.22	−2.09	4.87

Circle, circle drawing task; Tower, tower building; Fish, fish task; Animal, animal house; Reverse, reverse categorization; Bear, bear and dragon; DCCS, dimensional change card sort; D/N, day/night stroop.

relatively low, except for the *Animal House* task (for the 36–48 month sample), for which raw scores were transformed using an arcsine transformation, and the *DCCS* task, for which raw scores were transformed using a logarithmic transformation [Log10 (max range + 1 − x)]. The transformed descriptive statistics for the *Animal House* task and the *DCCS* task were as follows: *Animal House*: mean = 1.27, $SD = 0.38$, skewness = −1.06, kurtosis = 0.06; *DCCS*: mean = 0.85, $SD = 0.21$, skewness = 0.60, kurtosis = 0.99.

The mean scores obtained by both samples in the common inhibition measures were compared using an independent samples *t*-test. The results showed significantly better task performance for the older children in all tasks, including the *Circle Drawing Task* [$t_{(124)} = -3.58$, $p < 0.001$], the *Fish Task* [$t_{(125)} = -5.81$, $p < 0.001$] and the *Animal House* task [$t_{(127)} = -13.66$, $p < 0.001$].

CORRELATIONS

Zero-order (Pearson) and partial correlations controlled for age (upper triangle, **Table 2**) among inhibitory measures were performed.

Consistent with the findings of previous studies (Wiebe et al., 2011), the correlations were generally low in both samples.

In the 24–32 month sample, the *response inhibition* tasks were positively correlated with the *Fish Task*, which was considered to be an *interference suppression* task. In particular, the response to the incongruent condition of the *Fish Task* showed a significant correlation pattern with the slowdown motor response of the *Circle Drawing Task* and with the number of correct turns in the *Tower Building* task; these associations remained significant after controlling for age. Moreover, the *Tower Building* task showed a positive correlation with the *Reverse Categorization* task.

Among the *interference suppression* tasks, the number of correct items on the *Animal House* task correlated moderately with the correct responses on the *Fish Task* and with the number of correct items in the post-switch phase of the *Reverse Categorization* task. In this last case, the association was significant only after controlling for age.

In the 36–48 month sample, the *response inhibition* tasks correlated positively with one another. In particular, the ability to inhibit the interference to activate an alternative response of the *Day/Night Stroop* task was significantly correlated with the *Circle Drawing Task* and the ability to inhibit a prepotent response in the *Bear/Dragon* task. All of these tasks share the ability to inhibit an impulsive or a dominant response.

The *interference suppression* tasks were significantly correlated with one another. The *Animal House* task, which evaluates the ability to resolve a conflict generated by the previous presentation of a different classification rule, was positively correlated with the ability to manage interfering stimuli, as evaluated by the *Fish Task*, and with the ability to suppress the non-pertinent learned rule in a misleading situation of the *DCCS* task.

The *interference suppression* tasks were also correlated with *response inhibition* measures, such as the *Bear/Dragon* task and, in the case of the *Animal House* task and the *DCCS* task, with the *Day/Night Stroop* task.

CONFIRMATORY FACTOR ANALYSIS

To identify which model would be more useful to explain the observed data, a series of CFAs, based on covariance matrices, were performed using EQS 6.1 software[1] (Bentler, 2006). Multiple fit indices were considered for comparing models (for an example of an extensive description, see Schermelleh-Engel et al., 2003): the X^2 statistic, the root mean square error of approximation (RMSEA), the standardized root mean squared residual (SRMR), the Bentler's Comparative Fit Index (CFI), the Non-Normed Fit Index (NNFI) and the Akaike Information Criterion (AIC).

The X^2 test was used to evaluate the appropriateness of the CFA model: non-significant X^2 values indicated a minor difference between the covariance matrix generated by the model and the observed matrix and, thus, an acceptable fit. The RMSEA and the SRMR are the absolute fit indices, which assess how well an a priori model reproduces the sample data (Hu and Bentler, 1999). The RMSEA, which is a measure of the approximate fit in the population, measures how closely the covariances predicted by the model match the actual covariances. RMSEA values = 0.05 represent a good fit, values between 0.05 and 0.08 represent an adequate fit, values between 0.08 and

[1]EQS (6.1) [Computer software]. http://www.mvsoft.com/products.htm

Table 2 | Zero order and partial correlation controlled for age (upper triangle) between inhibitory measures in 24–32 month sample and in 36–48 month sample.

24–32 month sampleinhibition measures						36–48 month sampleinhibition measures						
	1	2	3	4	5		1	2	3	4	5	6
1. Circle	1	0.135	0.367**	0.082	0.168	1. Circle	1	0.011	0.281*	0.046	0.001	−0.044
2. Tower	0.068	1	0.241*	0.154	0.121	2. Bear	0.063	1	0.318**	0.362**	0.198	−0.258*
–	–	–	–	–	–	3. D/N	0.277*	0.360**	1	0.111	0.190	−0.171
3. Fish	0.351**	0.329*	1	0.325*	0.127	4. Fish	0.105	0.351**	0.133	1	0.291*	−0.141
4. Animal	0.100	0.119	0.329*	1	0.315*	5. Animal	0.098	0.255*	0.238*	0.328*	1	− 0.409**
5. Reverse	0.098	0.275*	0.178	0.189	1	6. DCCS	−0.017	−0.283*	−0.281*	−0.192	−0.399**	1

*Circle, circle drawing task; Tower, tower building; Fish, fish task; Animal, animal house; Reverse, reverse categorization; Bear, bear and dragon; DCCS, dimensional change card sort; D/N, day/night stroop. *$p < 0.05$; **$p < 0.01$. The negative values of the DCCS are due to the mathematical transformation used.*

0.10 represent a mediocre fit and values greater than 0.10 are not acceptable (Browne and Cudeck, 1993). The SRMR is the square root of the averaged squared residuals (i.e., the differences between the observed and predicted covariances). SRMR values < 0.10 are acceptable; however, values lower than 0.05 represent a good fit (Schermelleh-Engel et al., 2003). The CFI and the NNFI are incremental fit indices and measure the proportionate improvement in fit by comparing a target model with a baseline model (Hu and Bentler, 1999). The CFI compares the covariance matrix predicted by the model with the observed covariance matrix and compares the null model with the observed covariance matrix. The NNFI reflects the proportion by which the researcher's model improves the fit compared to the null model simultaneously controlling for the degrees of freedom. CFI and NNFI values greater than 0.97 are indicative of a good fit, whereas values greater than 0.95 may be interpreted as an acceptable fit (Schermelleh-Engel et al., 2003). The AIC statistic ($= X^2 - 2\ df$), which is a descriptive measure used to discriminate between competing models, was employed to compare the models. The models with the lowest AICs were considered to be the best.

Considering data separately from both age groups, two different theoretical models were performed: an inhibition unitary model and a two-factor model, in which *response inhibition* and *interference suppression* were distinguished. **Figure 1** schematically shows these comparative models. The fit indices for these models are summarized in **Table 3**.

For the 24–32 month sample, the unitary model was the only acceptable solution. The two-factor model results showed that the value of the correlation between the two dimensions was 1; thus, it was not possible to run a model in which the two dimensions are distinguished.

The unitary model showed no significant X^2 ($X^2 = 5.1$, $p = 0.40$) and acceptable to good fit indices. Specifically, the NNFI, the CFI and the RMSEA values indicated good fits, and the SRMR showed an acceptable fit.

In contrast, in the oldest sample, the two-factor model fits the data better than the more parsimonious single-factor model. Although the estimate correlation between factors is high ($r = 0.71$; the 95% confidence interval for the correlation was [0.50, 0.84]), the two-factor model allows a better explanation of the observed data. As presented in **Table 3**, X^2 was not significant in either solution (Model c, $X^2 = 11.27$, $p = 0.26$; Model d, $X^2 = 8.79$, $p = 0.36$). Nevertheless, the indices showed the best fit for the two-factor solution: the SRMR was acceptable in both models, however, in the two-factor model, the CFI, the NNFI and the RMSEA indicated good fits, whereas the same indices did not report acceptable values for the unitary model. Finally, the lowest AIC occurred for the two-factor model; thus, it showed the best fit.

As reported in **Figure 2**, in both age ranges, the models identified significantly predicted all observed variables (t values >2) with the exception of the *Reverse Categorization* task in the 24–32 month sample and the *Circle Drawing Task* in the 36–48 month sample. The proportion of variability explained by the tasks varied from 0.12 to 0.50 in the youngest children and from 0.21 to 0.42 in the oldest children.

DISCUSSION

The aim of the current study was to examine the nature of inhibitory processes in early childhood. Although several authors have suggested a multifaceted nature of inhibition (Nigg, 2000; Friedman and Miyake, 2004; Clark et al., 2013; Diamond, 2013), an empirical investigation of the latent organization of inhibitory processes in early childhood was missing. The present investigation was an initial attempt to empirically assess the fit of two alternative models, which describe the latent structure of inhibition during the period from toddlerhood to preschool, a key transition point in children's development during which substantial gains occur in inhibitory task performance (Kochanska et al., 1997; Diamond, 2002; Jones et al., 2003; Carlson, 2005; Garon et al., 2008).

Though the growth of inhibitory control during childhood has been largely documented, especially in toddlerhood and preschool years (Diamond, 2002, for a review), no confirmatory analysis had previously been conducted to investigate the latent structure of the cognitive processes involved in inhibition.

In the present study, two samples of children (ranging in age from 24 to 32 months and from 36 to 48 months) from various socio-demographic backgrounds were assessed using age-appropriate inhibitory tasks involving different response

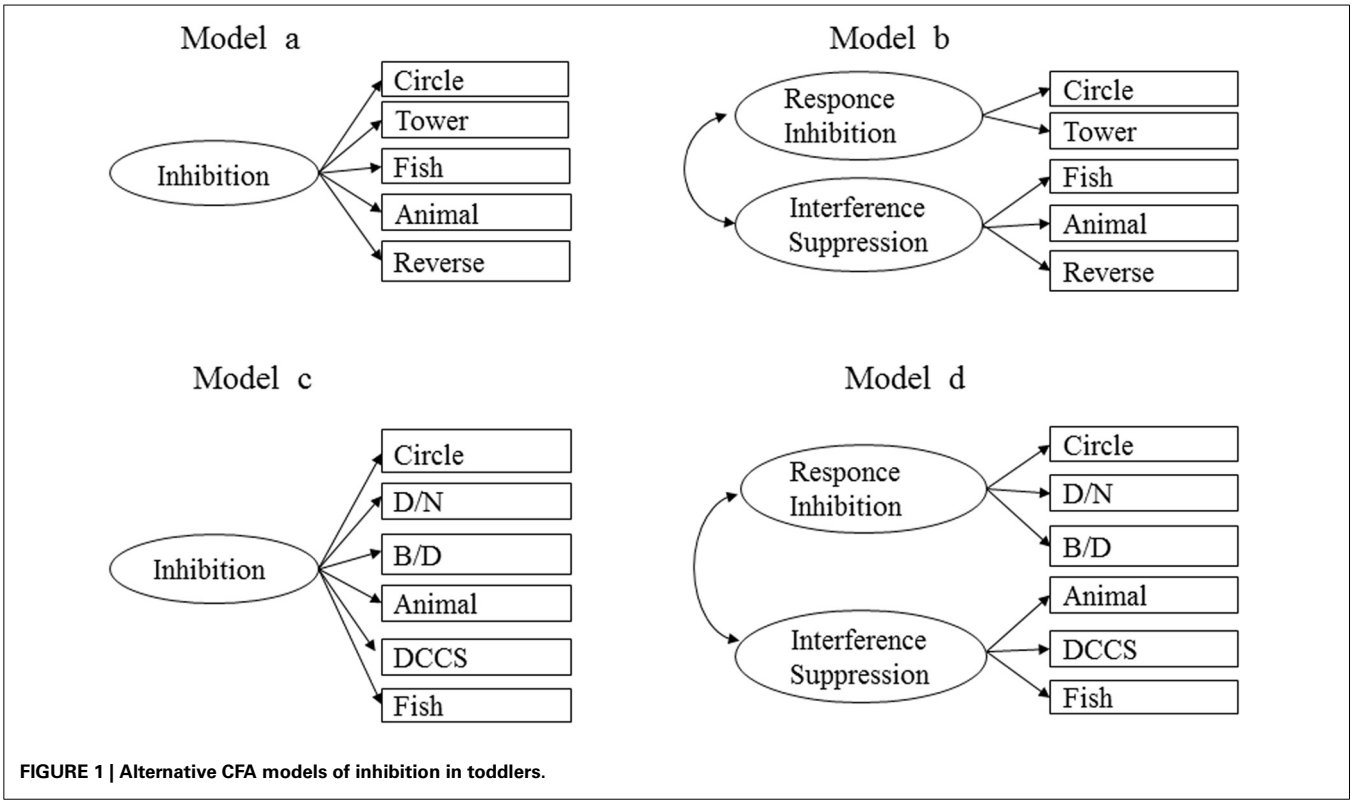


FIGURE 1 | Alternative CFA models of inhibition in toddlers.

Table 3 | Goodness of fit indices.

Group	Model		gdl	X^2	*p*	NNFI	CFI	RMSEA	SRMR	AIC
24–32 months	**Model a**	**One Factor**	5	5.100	0.40	0.985	0.993	0.020	0.062	−4.900
36–48 months	Model c	One Factor	9	11.270	0.26	0.898	0.939	0.061	0.066	−6.730
	Model d	**Two Factors**	8	8.786	0.36	0.960	0.979	0.038	0.058	−7.214

The endorsed models are indicated in bold.

RMSEA, root mean square error of approximation; SRMR, standardized root mean squared residual; CFI, comparative fit index; NNFI, non-normed fit index; AIC, akaike information criterion. The fit indices of Model b are not presented because the model did not converge.

demands. According to the literature, we considered two different models of inhibition development. First, we examined a unitary factor model, based on earlier studies that indicated a single undifferentiated executive control factor as the most appropriate for describing the executive latent structure in preschoolers (Wiebe et al., 2008, 2011; Hughes et al., 2010). Second, we examined a two-factor model, in which *response inhibition* was differentiated from *interference suppression*. The first component refers to the ability to control impulsive behavior and to prevent prepotent motor or verbal responses, whereas the second component involves more complex processes, such as WM, and comprises the suppression of interfering information. The separability of *response inhibition* and *interference suppression* has been described in older children (Bunge et al., 2002; Martin-Rhee and Bialystok, 2008).

In the 24–32 month sample, the simplest model with a single inhibition component was supported over the other model; the bi-factorial model, in which *response inhibition* and *interference* suppression were identified, was excluded because it was not acceptable due to the high correlation between the two latent factors. The unitary inhibition factor structure was chosen based on its relative and absolute model fits.

Parallel analyses were conducted for the 36–48 month sample. At this age level, the goodness-of-fit results indicated that a two-factor model provided the best fit to the data, with *response inhibition* as a separate dimension from an *interference suppression* factor.

The results suggest that inhibitory processes are not yet differentiated before 36 months of age, after which a distinction between different inhibitory dimensions emerges. We hypothesize a sequential development of inhibitory processes (see also Welsh et al., 1991; Espy, 1997; Espy et al., 2001; Senn et al., 2004): at an early age, the inhibition task performance primarily involves the ability to inhibit an impulsive or a dominant response (*response inhibition*); at a later stage, children develop a more cognitive inhibition that involves the suppression of interfering information or prepotent mental representations (*interference suppression*).

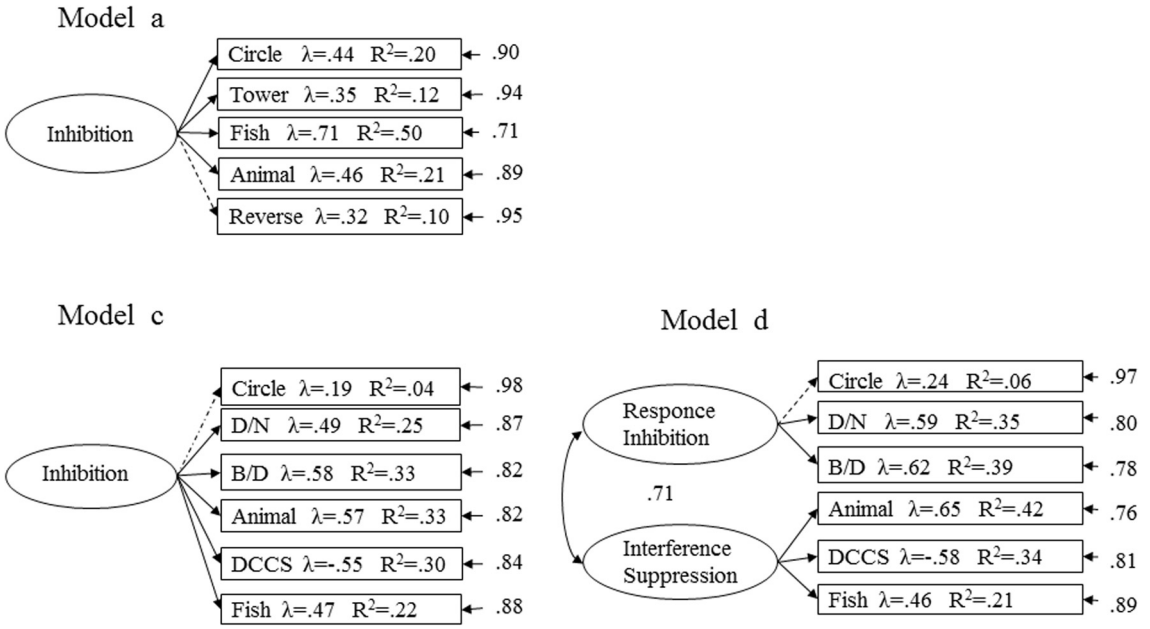


FIGURE 2 | Standardized factor loadings, R^2 values and error terms are shown for unitary and two factor models for both age groups: model a for youngest children and model c and model d for oldest children. The standardized factor loadings and the R^2 values are reported into the box for each observed variable. The error terms are shown near the observed variables at the end of the smaller, single-headed arrows. No significant factor loadings are shown with dotted arrows. The correlation between the factors for model d is that next to the curved, double-headed arrow. Note: Circle, circle drawing task; Tower, tower building; Fish, fish task; Animal, animal house; Reverse, reverse categorization; Bear, bear and dragon; DCCS, dimensional change card sort; D/N, day/night stroop.

Tasks were selected to maximize the difference between *response inhibition* and *interference suppression*. As Miller et al. (2012) suggested, the task selection and the choice of performance indicators may influence the findings; consequently, both factors must be selected to clearly separate the different cognitive processes that must be assessed.

Response inhibition was evaluated in toddlers using the *Circle Drawing Task* and the *Tower Building* task, which require the ability to suppress prepotent but inappropriate responses. Similarly, in the *Bear/Dragon* task, the child needs to selectively suppress commanded actions. In particular, he/she must choose between two conflicting response types (performing or suppressing an action) based on a rule; however, the child must respond to a single stimulus which is clearly indicated by the experimenter. In the *Day/Night Stroop* task, the child must also suppress the tendency to produce a dominant response in relation to a target. Both tasks are thought to require inhibitory control and WM in remembering the rules. However, memory demands do not significantly influence the performance. With regards to the *Day/Night Stroop* task, Gerstadt et al. (1994) demonstrated that if children are asked to associate the labels "day" and "night" to two abstract designs, even preschoolers succeed. This condition still requires remembering two rules, but it does not require inhibiting the tendency to say what the stimuli really represent (Diamond et al., 2002). In the *Bear/Dragon* task, Jones et al. (2003) found that children between 36 and 48 months of age performed accurately on the activation trials, with the percent of correct responses at all ages ranging from 90 to 94 percent; in contrast, the accuracy in the inhibition trials increased with age, suggesting that the main difficulty in this task is suppressing a prepotent response. Therefore, these tasks all have in common the request to suppress a response that is solicited by the stimulus; the go responses become prepotent because they are habitual (Simpson and Riggs, 2006). Reck and Hund (2011) found that the *Bear/Dragon* task and the *Day/Night Stroop* task loaded on the same factor in a sample of preschool children aged 3–6 years, which suggests that the two tasks assessed similar cognitive processes.

Interference suppression was evaluated using the *Fish* and the *Animal House* tasks in both toddlers and preschoolers, the *Reverse Categorization* task in toddlers, and the *DCCS* in preschoolers. These tasks all require some level of response inhibition, similar to the previous tasks; however, they also require an individual to filter out incongruent information within the stimuli because children must respond to stimuli that contain both relevant and distracting information.

In case of the *Fish Task,* children need to control the impulse to touch the stimulus before they have observed the fish's direction ("*I mustn't touch the fish's food immediately but I have to observe the fish's direction before*") (*response inhibition*); however, they also need to manage the visual and attentional interferences to solve the task. In particular, as suggested by Martin-Rhee and Bialystok (2008), this task requires the child to focus on one feature of the stimulus (the target fish direction) and ignore the other (the flankers' direction). This characteristic is present in all of the tasks that were chosen to assess *interference suppression.* The *Animal House* task requires children to control their impulsive behavior

of putting all the stickers in the colored house without following any rules. Nevertheless, each time, they also need to select the right piece of information that is necessary to accomplish the task; for example, to correctly place the duck, the child must select the blue house and ignore the houses with other colors.

The *Reverse Categorization* task, which was used in the 24–32 month sample, is a sorting task that is very similar to the *DCCS* task, which was administered to children aged 36–48 months. Both tasks require children to classify objects or cards by considering their different features, the blocks' size (big or small) in the case of the *Reverse Categorization* and the color (red or blue) or the shape (rabbit or boat) in the case of the *DCCS*. Toddlers often fail to classify the blocks in the *Reverse Categorization* task in the post-switch phase ("*Now I've to put the big block in the small box*") because they cannot inhibit the rule previously learned ("*I've to put the big block in the big box*"). Similarly, preschool children have difficulty switching from sorting by color to sorting by shape on the *DCCS* task because they have difficulty in inhibiting the old way of thinking about the objects. Children in the first year of preschool may remain stuck in thinking about objects according to the objects' initially relevant attribute (Diamond et al., 2005). The *DCCS* task requires high demands on the control of attention: children must inhibit their attention to a dimension that was previously valid to attend to a different aspect of the same stimulus.

The changes that occur in the nature of inhibitory processes from toddlerhood to early childhood may be due to both quantitative and qualitative changes in cognitive processing. For example, the *Fish* and *Animal House* tasks were explained by the same inhibitory dimension as the *Circle Drawing Task* in the 24–32 month sample. In contrast, in the 36–48 month sample, both tasks (i.e., the *Fish Task* and *Animal House* tasks) converged in the *interference suppression* factor, which suggests that at this age level, a child's performance is influenced by the ability to filter out irrelevant information, and not only by the ability to suppress a habitual response. Indeed, a specific task may not measure the same ability across different ages (Clark et al., 2013). For example, the *Tower of London* task (ToL), which is traditionally employed to assess planning in adults, proved to measure inhibitory control in young children (Bull et al., 2004).

As regards the reasons of the change in the organization of inhibitory processes across the two age-levels, this could be a result of maturational processes, as well as a consequence of the educational experiences of the children. While the 24–32 month sample was recruited from day-care centers, the children in the 36–48 month sample were attending the first year of preschool. In Italy, attendance at preschool is commonly accepted as the first essential stage of the educational system, and over 95% of children between 3 and 5 years of age attend a pre-primary school. Supporting school readiness, the preschool curriculum emphasizes activities that enhance creativity skills, social attitudes, autonomy and the learning process. The transition to preschool provides children with an opportunity to develop cognitive abilities and to improve self-regulation and executive function skills by increasing the children's participation in more structured activities that require more attentional control (Diamond and Lee, 2011; Hughes and Ensor, 2011).

However, the current results should be considered in the context of the study limitations.

First, because of the more limited behavioral repertoire of toddlers compared to preschool children, it was impossible to use exactly the same tasks in both age ranges. The tasks used in the 24–32 month sample are necessarily simpler, though they have similar inhibitory demands as the tasks used to assess the 36- to 48-month-old children. As indicated previously, the tasks used to assess *response inhibition* at both age levels comprised univalent stimuli associated with a prepotent response that must be overruled. While the tasks used to assess *interference suppression* comprised stimuli with different features, each was associated with a different response; thus, attention must be selectively focused on the relevant cue. Second, though the models tested were simple, the sample size at each age level was limited, suggesting that further evidence is needed to confirm our findings.

In conclusion, to the best of our knowledge, this is the first study to investigate the latent structure of inhibition at an early age. Because inhibition development is central in several theories of cognitive development (Dempster, 1992; Tipper, 1992; Harnishfeger and Bjorklund, 1993; Diamond and Taylor, 1996), the study of the nature of inhibitory processes from early childhood represents a significant area of research. Empirical evidence shows that children with typical development increase their performance in inhibition tasks from toddlerhood to the preschool period (Diamond, 2002; Carlson, 2005); at the same time, research has emphasized that a deficit in the development of inhibitory processes is associated with several psychopathological diseases, such as autism spectrum disorders (Ozonoff et al., 1991; Robinson et al., 2009) and attention deficit hyperactivity disorder (Barkley, 1997; Ozonoff and Jensen, 1999; Schachar et al., 2000). Finding an initial differentiation of inhibitory processes may be promising in understanding the development of inhibition in both typical and atypical developmental trajectories.

REFERENCES

Bachorowski, J. A., and Newman, J. P. (1985). Impulsivity in adults: motor inhibition and time-interval estimation. *Pers. Individ. Dif.* 6, 133–136. doi: 10.1016/0191-8869(85)90041-8

Barkley, R. A. (1997). *ADHD and the Nature of Self-control.* New York, NY: Guilford Press.

Bentler, P. M. (2006). *EQS 6 Structural Equations Program Manual.* Encino, CA: Multivariate Software.

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Blasi, G., Goldberg, T. E., Weickert, T., Das, S., Kohn, P., Zoltick, B., et al. (2006). Brain regions underlying response inhibition and interference monitoring and suppression. *Eur. J. Neurosci.* 23, 1658–1664. doi: 10.1111/j.1460-9568.2006.04680.x

Browne, M. W., and Cudeck, R. (1993). "Alternative ways of assessing model fit," in *Testing Structural Equation Models*, eds K. A. Bollen and J. S. Long (Beverly Hills, CA: Sage), 136–162.

Bull, R., Espy, K. A., and Senn, T. E. (2004). A comparison of performance on the Towers of London and Hanoi in young children. *J. Child Psychol. Psychiatry* 4, 743–754. doi: 10.1111/j.1469-7610.2004.00268.x

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., and Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron* 33, 301–311. doi: 10.1016/S0896-6273(01)00583-9

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carlson, S., Mandell, D., and Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Dev. Psychol.* 40, 1105–1122. doi: 10.1037/00121649.40.6.1105

Carlson, S. M., Moses, L. J., and Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Inf. Child Dev.* 11, 73–92. doi: 10.1002/icd.298

Clark, C. A. C., Sheffield, T. D., Chevalier, N., Nelson, J. M., Wiebe, S. A., and Espy, K. A. (2013). Charting early trajectories of executive control with the shape school. *Dev. Psychol.* 49, 1481–1493. doi: 10.1037/a0030578

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: toward a unified theory of cognitive development and aging. *Dev. Rev.* 12, 45–75. doi: 10.1016/0273-2297(92)90003-K

Diamond, A. (2002). "Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry," in *Principles of Frontal Lobe Function*, eds D. T. Stuss and R. T. Knight (New York, NY: Oxford University Press), 466–503. doi: 10.1093/acprof:oso/9780195134971.003.0029

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Diamond, A., Carlson, S. M., and Beck, D. M. (2005). Preschool children's performance in task switching on the dimensional change card sort task: separating the dimensions aids the ability to switch. *Dev. Neuropsychol.* 28, 689–729. doi: 10.1207/s15326942dn2802_7

Diamond, A., Kirkham, N., and Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Dev. Psychol.* 38, 352–362. doi: 10.1037/a0012593

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4-12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529

Diamond, A., and Taylor, C. (1996). Development of an aspect of executive control: development of the abilities to remember what I said and to "Do as I say, not as I do." *Dev. Psychobiol.* 29, 315–334. doi: 10.1002/(SICI)1098-2302(199605)29

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Espy, K. A. (1997). The shape school: assessing executive function in preschool children. *Dev. Neuropsychol.* 13, 495–499. doi: 10.1080/87565649709540690

Espy, K. A., Kaufmann, P. M., Glisky, M. L., and McDiarmid, M. D. (2001). New procedures to assess executive functions in preschool children. *Clin. Neuropsychol.* 15, 46–58. doi: 10.1076/clin.15.1.46.1908

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135. doi: 10.1037/0096-3445.133.1.101

Fuhs, M. W., and Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at Head Start. *Dev. Psychol.* 47, 404–416. doi: 10.1037/a0021065

Garon, N., Bryson, S., and Smith, I. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi: 10.1037/0033-2909.134.1.31

Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action: performance of children 3 ½-7 years old on a Stroop-like day-night test. *Cognition* 53, 129–153. doi: 10.1016/0010-0277(94)90068-X

Harnishfeger, K. K., and Bjorklund, D. F. (1993). "The ontogeny of inhibition mechanism: a renewed approach to cognitive development," in *Emerging Themes in Cognitive Development*, eds M. L. Howe and R. Pansak (New York, NY: Springer-Verlag), 28–49.

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: coventional criteria versus new alternatives. *Struct. Eq. Model.* 6, 1–55.doi: 10.1080/10705519909540118

Hughes, C., and Ensor, R. (2011). Individual differences in growth in Executive Function across the transition to school predict externalizing and internalizing behaviors and children's self-perceived academic success at age 6. *J. Exp. Child Psychol.* 108, 663–676. doi: 10.1016/j.jecp.2010.06.005

Hughes, C., Ensor, R., Wilson, A., and Graham, A. (2010). Tracking executive function across the transition to school: a latent variable approach. *Dev. Neuropsychol.* 35, 20–36. doi: 10.1080/87565640903325691

Jones, L. B., Rothbart, M. K., and Posner, M. I. (2003). Development of executive attention in preschool children. *Dev. Sci.* 6, 498–504. doi: 10.1111/1467-7687.00307

Kochanska, G., Murray, K., and Coy, C. K. (1997). Inhibitory control as a contributor to conscience in childhood: from toddler to early school age. *Child Dev.* 68, 263–277. doi: 10.2307/1131849

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., and Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Dev.* 67, 490–507. doi: 10.1111/j.1467-8624.1996.tb01747.x

Lehto, J. E., Juujärvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *British J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling Lang. Cogn.* 11, 81–93. doi: 10.1017/S1366728907003227

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., and Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *J. Cogn. Dev.* 13, 395–423. doi: 10.1080/15248372.2011.585478

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychol. Bull.* 126, 220–246. doi: 10.1037/0033-2909.126.2.220

Ozonoff, S., and Jensen, J. (1999). Specific executive function profiles in three neurodevelopmental disorders. *J. Autism Dev. Disord.* 29, 171–177. doi: 10.1023/A:1023052913110

Ozonoff, S., Pennington, B. F., and Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *J. Child Psychol. Psychiatry* 32, 1081–1105. doi: 10.1111/j.1469-7610.1991.tb00351.x

Reck, S. G., and Hund, A. M. (2011). Sustained attention and age predict inhibitory control during early childhood. *J. Exp. Child Psychol.* 108, 504–512. doi: 10.1016/j.jecp.2010.07.010

Reed, M., Pien, D. L., and Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill. Palmer. Q.* 30, 131–147.

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., and Howlin, P. (2009). Executive function in children with autism spectrum disorders. *Brain Cogn.* 71, 362–368. doi: 10. 1016/j.bandc.2009.06.007

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P. et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Rueda, M. R., Posner, M. I., and Rothbart, M. K. (2005). The development of executive attention: contributions to the emergence of self-regulation. *Dev. Neuropsychol.* 28, 573–594. doi: 10.1207/s15326942dn2802_2

Schachar, R., Mota, V. L., Logan, G. D., Tannock, R., and Klim, P. (2000). Confirmation of an inhibitory control deficit in attention-deficit/hyperactivity disorder. *J. Abnorm. Child Psychol.* 28, 227–235. doi: 10.1023/A:1005140103162

Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-ft measures. *Meth. Psychol. Res.* 8, 23–74. Available online at: http://www.mpr-online.de

Senn, T. E., Espy, K. A., and Kaufmann, P. M. (2004). Using path analysis to understand executive functions organization in preschool children. *Dev. Neuropsychol.* 26, 445–464. doi: 10.1207/s15326942dn2601_5

Simpson, A., and Riggs, K. J. (2006). Conditions under which children experience inhibitory difficulty with a "button-press" go/no-go task. *J. Exp. Child Psychol.* 94, 18–26. doi: 10.1016/j.jecp.2005.10.003

Simpson, A, Riggs, K. J, Beck, S. R, Gorniak, S. L, Wu, Y., Abbott, D., et al. (2012). Refining the understanding of inhibitory control: how response prepotency is created and overcome. *Dev. Sci.* 15, 62–73. doi: 10.1111/j.1467-7687.2011.01105.x

Tipper, S. P. (1992). Selections for actions: the role of inhibitory mechanisms. *Curr. Dir. Psychol. Sci.* 1, 105–112. doi: 10.1111/1467-8721.ep10768813

Usai, M. C., Viterbori, P., Traverso, L., and De Franchis, V. (2014). Latent structure of executive function in 5-to and 6-year-old children: a longitudinal study. *Eur. J. Dev. Psychol.* 11, 447–462. doi: 10.1080/17405629.2013.840578

Viterbori, P., Gandolfi, E., and Usai, M. C. (2012). Executive skills and early language development. *J. Appl. Psycholinguist.* 3, 17–32.

Wechsler, D. (1973). *Wechsler Preschool and Primary Scale of Intelligence—Italian Version*. Firenze: Organizzazioni Speciali.

Welsh, M. C., Pennington, B. F., and Groisser, D. B. (1991). A normative-developmental study of executive function: a window on prefrontal function in children. *Dev. Neuropsychol.* 7, 131–149. doi: 10.1080/87565649109540483

Wiebe, S., Espy, K., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575

Wiebe, S., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., and Espy, K. (2011). The structure of executive function in 3-year-olds. *J. Exp. Child Psychol.* 108, 436–452. doi: 10.1016/j.jecp.2010.08.008

Willoughby, M. T., Blair, C. B., Wirth, R. J., and Greenberg, M. (2010). The measurement of executive function at age 3 years: psychometric properties and criterion validity of a new battery of tasks. *Psychol. Assess.* 22, 306–317. doi: 10.1037/a0018708

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Prot.* 1, 297–301. doi: 10.1038/nprot.2006.46

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 January 2014; accepted: 11 April 2014; published online: 30 April 2014.
*Citation: Gandolfi E, Viterbori P, Traverso L and Usai MC (2014) Inhibitory processes in toddlers: a latent-variable approach. Front. Psychol. **5**:381. doi: 10.3389/fpsyg.2014.00381*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Gandolfi, Viterbori, Traverso and Usai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 22 July 2014
doi: 10.3389/fpsyg.2014.00733

Psychometric properties and convergent and predictive validity of an executive function test battery for two-year-olds

Hanna Mulder[1]*, Huub Hoofs[1,2], Josje Verhagen[1], Ineke van der Veen[3] and Paul P. M. Leseman[1]

[1] Department of Special Education, Faculty of Social Sciences, Utrecht University, Utrecht, Netherlands
[2] Department of Epidemiology, Faculty of Health Medicine and Life Sciences, CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht, Netherlands
[3] Kohnstamm Institute, University of Amsterdam, Amsterdam, Netherlands

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Stephanie Carlson, University of Minnesota, USA
Michael Willoughby, University of North Carolina at Chapel Hill, USA

***Correspondence:**
Hanna Mulder, Department of Special Education, Faculty of Social Sciences, Utrecht University, Martinus Langeveldgebouw, Heidelberglaan 1, PO Box 80.140, 3508 TC Utrecht, Netherlands
e-mail: h.mulder2@uu.nl

Executive function (EF) is an important predictor of numerous developmental outcomes, such as academic achievement and behavioral adjustment. Although a plethora of measurement instruments exists to assess executive function in children, only few of these are suitable for toddlers, and even fewer have undergone psychometric evaluation. The present study evaluates the psychometric properties and validity of an assessment battery for measuring EF in two-year-olds. A sample of 2437 children were administered the assessment battery at a mean age of 2;4 years ($SD = 0;3$ years) in a large-scale field study. Measures of both hot EF (snack and gift delay tasks) and cool EF (six boxes, memory for location, and visual search task) were included. Confirmatory Factor Analyses showed that a two-factor hot and cool EF model fitted the data better than a one-factor model. Measurement invariance was supported across groups differing in age, gender, socioeconomic status (SES), home language, and test setting. Criterion and convergent validity were evaluated by examining relationships between EF and age, gender, SES, home language, and parent and teacher reports of children's attention and inhibitory control. Predictive validity of the test battery was investigated by regressing children's pre-academic skills and behavioral problems at age three on the latent hot and cool EF factors at age 2 years. The test battery showed satisfactory psychometric quality and criterion, convergent, and predictive validity. Whereas cool EF predicted both pre-academic skills and behavior problems 1 year later, hot EF predicted behavior problems only. These results show that EF can be assessed with psychometrically sound instruments in children as young as 2 years, and that EF tasks can be reliably applied in large scale field research. The current instruments offer new opportunities for investigating EF in early childhood, and for evaluating interventions targeted at improving EF from a young age.

Keywords: executive function, toddlers, psychometrics, validity, delay of gratification, working memory, selective attention

INTRODUCTION

Executive function (EF) involves a wide array of cognitive processes needed for goal-directed behavior and self-regulation. In children and adults, EF has been shown to involve at least three main components: (i) working memory, defined as the ability to hold information in memory while performing mental operations on this information; (ii) inhibitory control, defined as the ability to suppress automatized and predominant responses; and (iii) shifting, or the ability to change cognitive set in order to switch between different tasks (Miyake et al., 2000; Davidson et al., 2006; Garon et al., 2008). There is growing evidence that EF is a strong predictor of various aspects of child development, such as academic skills. Specifically, studies have found that EF ability at preschool age predicts later academic achievement (Blair and Razza, 2007; Clark et al., 2010). Moreover, development of EF over the preschool years, or growth in EF, is related to growth in academic skills such as math, vocabulary and emergent literacy (Mcclelland et al., 2007; Raver et al., 2011; Van der Ven et al., 2012). Finally, EF is important for more general learning-related skills, such as work attitude (Blair et al., 2005; Ponitz et al., 2009), and socioemotional skills (Denham et al., 2012). Given the importance of EF at a young age for later academic and behavioral functioning, there is a clear need for valid and psychometrically sound instruments to assess EF in early childhood. To date, however, few EF tasks are available for use with children younger than 3 years of age, and the instruments that are available most often have not been evaluated psychometrically. Such a psychometric evaluation is crucial as "the results will only be as good as the test," which entails that only valid and reliable assessment tools will contribute to our understanding of young children's EF and

thereby help to prevent academic failure from a young age (Blair and Diamond, 2008).

Although many studies have investigated EF in preschoolers aged between 3 and 5 years in the past years (Wiebe et al., 2008, 2011; Willoughby et al., 2010) not much is known about EF development in toddlers (cf. Garon et al., 2008). In particular, two-year-olds are a neglected group in research on EF development. Rose et al. (2009) noted that there is a gap in our knowledge about cognitive development in toddlerhood, and others even have described the period between 2 and 3 years of life as the "dark ages" of cognitive development (Meltzoff et al., 1999). One of the reasons for this gap in the literature is undoubtedly the relative difficulty of testing toddlers (see also Hughes and Ensor, 2005). Children this young generally have short attention spans, limited motor skills, and they do not yet dispose of complex language skills. As such, EF measures designed for preschoolers tend to be too challenging for toddlers. Thus, in order to assess rather complex processes such as controlling a dominant response or updating information in memory, tasks have to be developed that measure these abilities while not burdening children's motor, attentional and linguistic skills too much.

For two-year-old children, a few studies have looked at (the development of) EF and/or the relationships with other developmental domains such as theory of mind (Carlson et al., 2004; Hughes and Ensor, 2005, 2007; Miller and Marcovitch, 2011; Fitzpatrick and Pagani, 2012). With some exceptions (Hughes and Ensor, 2005, 2007; Fitzpatrick and Pagani, 2012), most studies have included relatively small samples of children that were tested in highly controlled laboratory settings. Consequently, there often is a high overrepresentation of children from motivated, high socioeconomic status (SES) parents willing to participate in a study, which seriously limits most studies' external validity (see also Willoughby et al., 2010). Also, a close psychometric scrutiny of the EF assessments used in these studies is generally absent.

An exception to this is a study by Carlson (2005), who addressed the psychometric properties of EF assessment tools in two- to six-year-olds, including a sample of 118 two-year-olds. This study showed relatively strong discriminatory power for most tasks for toddlers, enabling a proper differentiation between children of varying EF ability. The sample consisted of children from predominantly middle-class Caucasian families, however. Likewise, Garon et al. (2013) evaluated a battery of tasks assessing working memory, inhibition, and shifting for children aged 18–67 months. This study showed that the EF battery was sensitive to developmental improvements across this age span, and internal consistency of each of the measures was adequate to good. Again, however, the sample contained mostly middle-class families, leaving unanswered the question as to how appropriate such measures are for children from different socio-economic and ethnic backgrounds. Thus, although a few previous studies have assessed the psychometric properties of EF measures in samples with toddlers, these studies included mostly relatively high SES families, leaving unclear how appropriate such tasks are for children from less advantaged backgrounds.

The current study adds to the available literature on the psychometric quality of EF tasks for young children by investigating the psychometric properties of a battery of EF measures in a large sample of two-year-olds from diverse socio-economic and ethnic backgrounds. The EF battery in our study was designed for the purposes of evaluating effects of preschool education and care in the Netherlands on later socio-emotional, cognitive and academic skills (cf. Pre-COOL, see below). Therefore, we aimed to include a set of measures which would predict child developmental outcomes across multiple domains. Our decisions were informed by the literature about the development and factor structure of EF in young children as well as the predictive value of EF tasks for future academic and socio-emotional skills. The factor structure of EF has typically been investigated used Confirmatory Factor Analysis (CFA), a statistical approach which allows for modeling of shared variance amongst constructs. Through using CFA, conclusions can be drawn about the way different tasks cluster together, providing information about the underlying, latent factors which drive task performance rather than specific tasks. In the next sections, we describe two different lines of research on EF in child development, their main findings on the organizational structure of EF in young children, and the predictive validity of the EF construct(s) for developmental outcomes that guided our decisions when designing our task battery.

Miyake et al. (2000) have shown that the structure of EF comprises separate but interrelated inhibition, shifting, and working memory factors in adults. In a recent revision of their theory, however, Miyake and Friedman (2012) showed that EF in adults is best represented by a common EF factor and separate updating-specific and shifting-specific factors. So, their previous inhibition factor now fully overlaps with the common EF factor in this account. Although differentiation of EF into the three components of inhibition, shifting, and working memory in children has been confirmed by Lehto et al. (2003), a number of other studies support a two-factor over a three-factor model in childhood (Van der Ven et al., 2013; Usai et al., 2014). Van der Ven et al. (2013) argued that measurement selection, which varies widely across the EF literature, may be at the core of the variation in findings between studies. A similar pattern of findings occurs in the preschool EF literature. While some previous studies have found a single latent EF construct in preschoolers (Wiebe et al., 2008, 2011), others have observed more differentiated EF skills already at this young age (Garon et al., 2013). Miller et al. (2012) studied the factor structure of EF in three- to five-year-old children using working memory, inhibitory control, and shifting measures. In their first set of analyses, they replicated the finding by Wiebe et al. (2008) that EF comprises a single latent factor at this young age. However, in a second set of analyses, they selected different response indicators for some of their measures and found that a two-factor model with separate but related working memory and inhibition factors fitted their data better than a single or three-factor model. Miller et al. (2012) concluded that measurement and response indicator selection is crucial and may explain different findings across studies, in line with the claims made by Van der Ven et al. (2013). More clear-cut evidence regarding the role of age in the development of the structure of EF across development comes from studies which have administered the same EF battery to children of different ages and investigated measurement invariance across age. For example, Wiebe et al. (2008) found that a unitary EF model fitted their data best in a study of 2.3- to 6-year old

children including a comprehensive battery of inhibitory control and working memory measures. They found that their measurement model was invariant across age, indicating that a unitary factor fitted the data well for both younger and older preschoolers. Moreover, Shing et al. (2010) studied a battery of inhibitory control and working memory tasks in children aged 4–14.5 years old. They observed increasing fractionation of EF with age; a single latent factor was observed for their two youngest age groups, while separate working memory and inhibitory control factors were observed in their oldest age group. Thus, although differences in tasks and/or response selection may to a large extent explain differences in findings between studies regarding the fractionation of EF in preschoolers and older children, there is some evidence that EF is a unitary factor in preschool children and only becomes more fractionated when children grow older.

In a separate line of research, a distinction has been made between executive processing of neutral cognitive and affective stimuli. The former typically involve measures of inhibition, shifting, and working memory as described above. The latter is most often limited to assessments of inhibitory control in the face of an affective stimulus, and is usually assessed with delay of gratification tasks, which require the child to suppress touching an attractive object or sweet (Kochanska et al., 2000). Confirmatory factor analyses in studies of young children have shown that the executive processing of cognitive and affective stimuli are typically represented by separate latent factors, labeled "cool" and "hot" EF, respectively (Brock et al., 2009; Willoughby et al., 2011; Bassett et al., 2012). Again, however, there is a discrepancy between studies. While most studies have found that a two-factor model with separate hot and cool factors fitted the data best (Brock et al., 2009; Willoughby et al., 2011; Bassett et al., 2012), others have found that a two-factor model does not fit the data better than a single EF factor model in preschoolers (Allan and Lonigan, 2011). However, investigations of the predictive validity of cool and hot EF as separate factors lend support to their differentiation. In a study by Willoughby et al. (2011), cool EF was predictive of academic performance, while hot EF was predictive of behavioral adjustment in preschoolers. Similarly, Kim et al. (2013) showed that latent cool and hot EF factors differentially predicted academic skills and behavior problems: Cool EF predicted academic performance, while hot EF predicted behavior problems. In contrast to the studies by Willoughby et al. (2011) and Kim et al. (2013), Brock et al. (2009) found that cool EF predicted learning-related behaviors, classroom engagement, and math skills in kindergarteners, while hot EF predicted none of these outcomes when analyzed concurrently with cool EF. Thus, although the differentiation between hot and cool EF is not always confirmed and the theoretical debate about the meaning of this distinction is still ongoing (Welsh and Peterson, 2014), there are clear indications that hot and cool EF measures differentially predict developmental outcomes. An important question that remains is whether hot and cool EF can be distinguished already before preschool age and whether they are differentially predictive of developmental outcomes at this young age, given that all previous studies are on older children.

Since hot and cool EF measures may be differentially predictive of academic and behavioral outcomes, we included measures of hot as well as cool EF in our task battery for toddlers. For each domain, we selected multiple measures, to be able to use CFA and work with latent factors. As argued above, the main advantage of this approach is that task-specific measurement error can be partialled out from the latent constructs under investigation, which is especially beneficial in studies on young children, where measurement error tends to be large. For example, Willoughby et al. (2010) have shown that the association between EF measures and parent, teacher, and research assistant ratings of hyperactivity in three-year olds were weak to moderate for separate test measures, but when EF was modeled as a latent factor, the association with informant ratings of hyperactivity became much stronger. Willoughby et al. conclude that, as the separate measures are confounded with measurement error and task-specific demands (e.g., motor or verbal skills) and the latent factor represents only shared variance across measures, the latent measures provides a more reliable estimate of EF ability. Further evidence comes from test-retest analyses of an EF battery for four-year-olds, showing that test-retest reliability was much higher for a latent EF ability construct than for each separate measure alone (Willoughby and Blair, 2011). Based on these findings, we decided to include multiple measures of the hot and cool constructs in our EF battery. In the cool EF domain, we included two tasks assessing working memory and one task measuring selective attention. We initially also included an inhibitory control task (an adaptation of the Shapes task, Kochanska et al., 2000), but this task proved to be too difficult for the younger children in our sample and was dropped from the battery. In the hot EF domain, we included two delay of gratification tasks, a snack and gift delay (Kochanska et al., 2000).

AIMS OF THE CURRENT STUDY

The aims of our study were twofold: (1) investigate the psychometric properties of our EF task battery for toddlers, and (2) study criterion, convergent, and predictive validity of the test battery. To evaluate the psychometric properties of the test battery, the following steps were taken. First, we applied CFA to evaluate a two-factor hot and cool measurement model and compare this model to a one-factor model. Based on the previous studies described above, we expected to find support for the two-factor over the one-factor model. Second, as children in our sample were either tested at their daycare center or at home, and comprised a mixed group in terms of their language background and SES, we studied whether our measurement model was invariant across a number of groups: SES (low/middle vs. high SES), age ($<$2.5 years vs. $>$2.5 years of age), assessment setting (home vs. day care center), home language (monolingual Dutch vs. non-monolingual Dutch), and gender. If measurement model invariance is supported across groups, this implies that measures relate to the latent constructs in the same way across different groups, allowing for a fair comparison between different subgroups of children.

Criterion validity was studied by examining relationships between children's latent EF abilities and gender, SES, home language, age, and assessment setting. Previous studies investigating gender differences in EF have yielded mixed results, with some studies showing that girls outperformed boys (Kochanska and Knaack, 2003; Wiebe et al., 2008), and others showing no gender differences (Wiebe et al., 2011). Two recent studies have

investigated gender differences in EF in children across different cultural contexts, using the Head-Toes-Knees-Shoulders task (Ponitz et al., 2009). Gender differences were observed in the United States and Iceland, but not in Taiwan, China, South Korea, Germany, and France, suggesting that cultural differences in socialization practices might play a role in the emergence of gender differences in EF (Wanless et al., 2013; Gestsdottir et al., 2014). Not much is known about gender differences in EF in children in the Netherlands, although Huizinga and Smidts (2011) found that Dutch five to 18-year-old girls received higher ratings on EF by their parents than boys. Therefore, we expected that, if a gender difference was observed, girls would perform better than boys. Furthermore, children from lower SES families were expected to obtain lower scores than children from high SES families (Hughes and Ensor, 2005; Noble et al., 2005, 2007). As for home language, no clear prediction could be formulated. Previous studies have shown that bilingual children may show enhanced EF as compared to their monolingual peers, already at preschool age, but in young children, this EF advantage seems to be restricted to native bilingual children who are exposed to two languages at home from birth (Blom et al., in press; Carlson and Meltzoff, 2008; Poulin-Dubois et al., 2011). In our sample, a large number of children were predominantly exposed to another language than Dutch at home. These children may score lower on the EF tasks which were administered in Dutch, due to their poorer knowledge of the Dutch language. In addition, associated with a different home language, different cultural customs regarding early play and cognitive stimulation can be at stake that can influence EF scores. As for age, previous studies have shown significant growth in EF skills during the third year of life (Garon et al., 2013). Therefore, we expected a strong effect of age on EF ability.

Convergent validity of the test battery was assessed through studying the association between children's EF ability and parent and teacher reports of children's attention and inhibitory control. Parent- and teacher-rated attentional focusing and inhibitory control scores were expected to be positively related to children's EF scores, as these two temperament dimensions are conceptually related to EF (Rothbart et al., 2003; Blair and Razza, 2007). Finally, predictive validity was assessed by regressing children's pre-academic skills and behavior problems at preschool age on children's EF scores at toddler age. Whereas hot EF was expected to predict behavior problems, cool EF was expected to predict pre-academic skills (Willoughby et al., 2011; Kim et al., 2013).

METHODS

PARTICIPANTS

Children participating in this study were involved in the longitudinal national cohort study Pre-COOL on the effectiveness of preschool care and educational provisions in the Netherlands, commissioned by the Dutch Ministry of Education, Culture and Sciences (Veen et al., 2012). In Pre-COOL, children are being assessed longitudinally from age 2–5 years. At the first wave of assessment, children were aged 2 years (M = 2;4 years, SD = 0;3 years, range = 1;8–3;1 years). Although the age range was wide, 70% of children were aged between 2;0–2;6 years, 28% were aged between 2;6 and 3;0 years, only 22 children were below 2;0 years (<0.01%), and six children were older than 3;0 years (<0.01%). At the second measurement wave, children were aged 3;6 years on average (SD = 0;2; range 2;11–4;5 years). The average time interval between assessments was 1;2 years (SD = 0;4; range 0;6 to 2;2 years). Gender was equally distributed (49% girls). As for SES, 41.5% of the children were from low/middle SES families and 58.5% came from high SES families. Most children were from monolingual Dutch homes; 28% of the children came non-monolingual Dutch families. The sample consisted of two sub-samples: a center-based sample which included children participating in center-based education and care and recruited through their center, and a home-based sample which included children recruited through the municipal registration records (and as such includes both children attending day care and children not attending day care). The sample was geographically well-spread across rural, semi-urban, and urban areas in all parts of the Netherlands. Approval for the study was obtained from the Ethical Advisory Committee of the Faculty of Social and Behavioral Sciences of Utrecht University[1].

Center-based sample

The Pre-COOL study is linked to the national cohort study COOL. The latter is aimed at following students' educational careers in Dutch primary and secondary education from age 5 to 18 years. Children in the Pre-COOL sample will enroll in the COOL study, so they can be followed from toddler age through to late adolescence. To increase the likelihood of Pre-COOL participants entering primary schools involved in COOL, recruitment of the center-based sample proceeded in a number of steps. First, primary schools participating in COOL were selected. 300 primary schools, randomly drawn from the COOL cohort, were approached. 139 schools agreed to participate. Next, COOL primary schools were asked to identify the preschool day care and education centers that were attended by most of their new students. In addition, municipal records and the internet were used to identify additional preschool care and education centers in the same neighborhoods as the COOL schools. Over 500 centers across the Netherlands were invited to participate in Pre-COOL, of which 289 centers agreed to take part. Finally, children born between April 1 and November 1, 2008, were identified in these centers. Parents of eligible children were personally informed by their child's teacher about the Pre-COOL study and were given a letter containing information about the study, explicitly giving them the opportunity to withdraw their child from participation by notifying the teacher. In total, 1819 children enrolled in the center-based sample.

Home-based sample

A sample of 6000 families with a child born between April 1 and November 1, 2008, living in neighborhoods close to the participating COOL schools was drawn from the municipal

[1]The research reported in this article involves healthy human participants, and does not utilize any invasive techniques, substance administration or psychological manipulations. Therefore, compliant with Dutch law, this study only required, and received approval from our internal faculty board (Faculty's Advisory Committee under the Medical Research (Human Subjects) Act (WMO Advisory Committee) at Utrecht University.

population registers. Parents received a letter in which they were invited to take part in the study with a pre-paid answering card. Additionally, families with an immigration background living in Pre-COOL neighborhoods in the urban agglomerations of Amsterdam, Rotterdam and The Hague were contacted personally during home visits in order to increase participation from these groups. In total, 1139 parents responded to the study invitation. Of those, 1008 agreed to participate in the study.

PROCEDURES

Children participating in the study were assessed at home (home-based sample) or at their center (center-based sample) at both waves. Testing took place in a quiet room. The tests in this study were part of a more comprehensive test battery which took on average 45 min to administer. Tests were given in a fixed order. At the first wave, two computerized language tasks, the visual search task, two further computerized language tasks, the snack delay, memory for location, six boxes, and gift delay task were given. At the second wave, a computerized language task, the vocabulary task, visual search task, two further computerized language tasks, and a computerized EF task, gift delay, emergent math, six boxes, and a second delay task were given. Research assistants (RA's) allowed children to have short breaks when necessary. Parents and teachers were asked to fill out a questionnaire with items addressing, among others, demographic variables and children's temperament and behavior.

To secure standardized assessment, RA's went through an intensive training phase before they were allowed to start data collection in each study wave. First, they attended a full day test administration course. Second, they received a very detailed standardized test protocol with step-by-step descriptions of the procedures for each measure. Third, they submitted a video recording of a practice session with a two-year-old to the study center, together with their scoring forms. The test administration procedures and scoring forms were carefully reviewed by the first and third author, and each RA was sent a detailed feedback report. This report was discussed by telephone. If the RA followed the standardized protocol, they were allowed to start data collection. If major administration or coding errors were observed, the RA was required to submit a second video for feedback purposes. The first and third author discussed any difficult cases until agreement was reached, and read each other's feedback reports before sending them to RA's, to ensure that no divergence in their evaluations occurred throughout the process.

MEASURES

At the first wave, children completed the EF tasks, parents rated children's inhibitory control and attentional focusing, and teachers rated children's inhibitory control and work attitude in the classroom. At the second wave, children's emergent math skills and vocabulary were assessed, and parents and teachers rated children's externalizing behavior problems. Each of the measures is described below in turn. It should be noted that teacher ratings were only available for the center-based sample.

Wave one measures

Attention (visual search). To measure selective attention, a computerized visual search task was developed for the purposes of the present study, based on the work by Gerhardstein and Rovee-Collier (2002), and Scerif et al. (2004). In this task, children were shown a structured display of 48 animals on a 6 × 8 grid on the laptop screen using E-Prime 2.0 (Schneider et al., 2002). Stimuli were images of elephants, bears, and donkeys, which were the same in color and size. Children were instructed to locate as many targets (elephants) as possible while ignoring the distractors (bears and donkeys). As such, children had to try to focus their attention only on the targets while suppressing interfering visual stimuli. To minimize memory demands, the targets that the child had located were crossed off with a line by the assessor. Following three practice trials, children were given three test items which lasted 40 s each. Each test item contained eight targets. Throughout the test items, children were encouraged to search as fast as possible and were continuously given feedback according to protocol (i.e., when the child pointed to a target: "Well done! Can you find another elephant?" or when the child pointed to a distractor: "No, where is an elephant?" or when the child pointed to the same elephant twice: "No, where is another elephant?"). Feedback rules were developed following careful piloting. Corrective feedback was used to ensure memory demands of this task were minimal. Accuracy for each test item was scored and averaged across items (i.e., the number of targets located correctly within the time limit, range 0–8). When children achieved a total score of "0", indicating that they did not find any targets on the three test items, their score was set to "missing," as we cannot be completely certain that they understood the task rules properly.

Visuospatial working memory task (six boxes). The six boxes task (Diamond et al., 1997) was used to measure visuospatial working memory capacity. To familiarize children with the task, a practice trial was given in which the child was shown how two wooden toys were hidden in two identical white boxes with blue lids. The child was then instructed to retrieve the toys one by one. The RA distracted the child for 1 s in between the two search attempts. If the child failed the practice trial (i.e., the child didn't find both toys in two search attempts), this procedure was repeated. After the practice trials, the test trials were given.

For the six test trials, six different wooden toys were hidden in six identical white boxes with blue lids while the child watched. The boxes were placed in two slightly asymmetrical rows of three boxes, rather than two perfectly aligned rows, to discourage the use of a simple strategy of opening the boxes row by row. Children were given six search attempts to find all toys. They were actively distracted by the RA for 6 s in between search attempts, as pilot work had shown that a 6 s delay gave the most optimal distribution of scores for this age range. After the child had taken a toy out of a box, the RA showed them clearly that that box was empty before closing the lid again ("Look, this one is empty now!"). If children moved a box without opening it (for example, by shaking it lightly to hear if it contained a toy), the RA opened the box and this box was scored as the child's choice for that search attempt. On both the practice and test trials, children were given positive feedback when they opened a box containing a toy. However, when they opened a box that was already empty, they were told "Oh no, that one is empty" to encourage them to search in a different box at the next search attempt. Thus, in

this task, children had to try to remember which boxes they had already emptied and which boxes still contained a toy and retain this information over the delay time. Accuracy across test trials (i.e., the number of toys obtained correctly) was scored for each child.

Visuospatial short-term memory span task (memory for location). This task assesses visuospatial memory span and was based on work by Pelphrey et al. (2004) and Vicari et al. (2004). The procedure of this task was similar to that of the six boxes task: Children were shown how a different set of small wooden figures was hidden in six identical white boxes which were placed in two symmetrical rows of three boxes each. However, in contrast to the six boxes task, the number of figures hidden varied across test items (range 1–4). After hiding the figures, the RA distracted the child for 1 s, and the child was then asked to find all the figures for that item. An item was scored as correct if the child retrieved all hidden figures in the minimum number of search attempts.

For this task, an adaptive testing procedure was used in which task difficulty level increased after each successful item. Difficulty level was defined as the number of hidden figures and ranged from one to four. This difficulty level was based on previous work showing that 24-month old toddlers were able to hold between two and three items in memory (Rose et al., 2009), and our own pilot work with children between age 2 and 3 years.

On the first test item, one figure was hidden. If the child passed this item, difficulty level was increased, and two figures were hidden on the next item. However, if the child failed the first item, an additional item with one figure was given. Children received up to two trials for each difficulty level, with the exception of the first level for which children received up to three trials to familiarize them with the procedure. If children failed all items at a given difficulty level, task administration was discontinued. Throughout the task, children were given feedback in a similar fashion as during the six boxes task ("Well done!" when they found a toy and "Oh no, that one is empty" when opening a box which did not contain a toy). The number of locations that a child could retain in memory simultaneously was measured in this task. Scores were calculated as the highest level (i.e., span) performed correctly for each child (range 0–4).

Delay of gratification (snack delay). The snack delay task was a simplified version of the Kochanska et al. (2000) snack delay task. In this task, an open box of raisins was placed in front of the child on the table at a distance of 25 cm. The child was then instructed to try not to touch the box of raisins until the RA had finished another task. The RA then moved away out of sight of the child and observed the child's behavior for 1 min. After the delay time, the child was always given positive feedback and they were given the box of raisins (if they had not already taken the box themselves). Three different behaviors were coded by the RA during the delay time: (1) touching the box or raisins, (2) picking up the box or raisins, and (3) eating the raisins. The occurrence of each behavior was coded as present (0) or absent (1) during the delay, so that a higher score indicated better task performance. The sum across these behavioral codes was scored (range 0–3). Children who obtained a total score of 1 or 2 were collapsed into one group due to a low number of children obtaining these scores (i.e., most children either ate the raisins or refrained from touching them). The total score then ranged from 0–2.

Delay of gratification (gift delay). The gift delay task was an adaptation of the Kochanska et al. (2000) gift delay task. This task was similar to the snack delay task, except that the box of raisins was replaced by an attractively wrapped gift with a bow. The child was instructed to try not to touch the gift during a delay of 1 min. The occurrence of three different behaviors was coded by the RA during the delay time: (1) touching the gift or bow, (2) tearing the wrapping paper, and (3) unpacking the gift completely (i.e., by taking the gift, a small rubber duck, out of the wrapping paper). However, the third category, unwrapping the gift completely, turned out to be too demanding for the motor skills of children this young, and was omitted from the analyses. The occurrence of each of the remaining two behaviors was coded as present (0) or absent (1) during the delay time of 1 min, so that a higher score indicated better task performance. The total score for this task was the sum across these behavioral codes (range 0–2).

In a separate study, video observations of the snack and gift delay tasks were coded to determine the reliability of the live codes in a sample of Dutch two- and three-year-olds. Kappa's were as follows for the snack delay task ($N = 59$): 0.96 for touching behavior and picking up the box of raisins combined, and 0.90 for eating the raisins. Agreement between video and live codes was 98.3 and 96.6%, respectively (chance level of agreement: 50%). For the gift delay task, the following Kappa's were observed ($N = 53$): 0.89 for touching behavior, and 0.74 for tearing the wrapping paper. Agreement between video and live codes was 96.2 and 94.3%, respectively (chance level of agreement: 50%).

Parent and teacher ratings of inhibitory control and parent ratings of attentional focusing (Early Childhood Behavior Questionnaire). The parent and teacher rated constructs inhibitory control and attentional focusing were assessed using a shortened version of the Dutch version of the Early Childhood Behavior Questionnaire (ECBQ, Putnam et al., 2006). This questionnaire was filled out by children's parents (six items for inhibitory control, four items for attentional focusing) and one of their teachers (three items for inhibitory control). As participating children in the center sample were often in the same group, many teachers had to fill out the questionnaire for more than one child in their group. Thus, very few items were selected for use with teachers to keep the questionnaire as short as possible. Items were selected based on pilot work with 56 parents and 44 teachers of two- to three-year-olds. Although the ECBQ is originally designed for use with parents (Putnam et al., 2006), we made minimal adaptations to questionnaire items for use with daycare teachers (i.e., "your child" in the parent questionnaire was replaced by "this child" in the teacher questionnaire'). For each item, respondents were asked to indicate the frequency with which a certain behavior (e.g., "ignoring a warning") occurred on a seven-point Likert scale (from "never" to "always"). Example items are: "When told no, how often did your/this child ignore your warning?" (inhibitory control) and

"When engaged in an activity requiring attention, such as building with blocks, how often did your child stay involved for 10 min or more?" (attentional focusing). Cronbach's alpha's were 0.78 and 0.84 for parent and teacher rated inhibitory control, respectively, and 0.78 for parent rated attentional focusing.

Teacher ratings of children's attention. Teachers of children in the center cohort also reported on children's attention during play and work using a four-item scale based on a short questionnaire designed for the COOL study (Driessen et al., 2009) and the SCHOBL-R (Bleichrodt et al., 1993). This tool has been found to be appropriate for collecting data on children's behavior in center-based care and education settings. Items concerning classroom behaviors (e.g., "works carefully," "is attentive") were rated on a five-point Likert scale (from "definitely untrue" to "definitely true"). Cronbach's alpha was 0.80.

Wave two measures

Parent and teacher ratings of children's externalizing behavior problems. To assess children's externalizing problem behavior, caregivers were asked to rate five items of the Problem Scale of the Brief Infant-Toddler Social and Emotional Assessment (BITSEA) (Briggs-Gowan and Carter, 2001). The following aspects of externalizing problem behavior are included in the BITSEA: activity/impulsivity, aggression/defiance, and peer aggression. The selection of items from the original Problem Scale was based on pilot data. Criteria were: inclusion of all three topics, discriminatory power, good internal consistency, and suitability of items for both parents and caregivers. Example items are: "Is very loud" (activity/impulsivity), "Purposely tries to hurt you (or other parent)" (aggression/defiance), and "Hits, shoves, kicks, or bites children (not including brother/sister)" (peer aggression). Cronbach's alpha was 0.85 for teachers and 0.86 for parents.

Children's emerging math skills. Children's emergent math skills were assessed with a short version of the Math Test for Toddlers developed by the Dutch National Institute for Educational Measurement (CITO) (Op den Kamp and Keuning, 2011). About two thirds of the total number of test items (15) were selected by CITO, based on suitability of difficulty, discriminatory power and adequacy of reliability (of 0.70). To gain a more even distribution of items across topics/aspects, one item was added. The final selection of 16 items covered three aspects: number sense, measurement, and geometry. Using item response theory (IRT) modeling, a skill score was calculated by CITO based on the responses to the 16 items.

Children's vocabulary. Receptive vocabulary was assessed with the Dutch version of the Peabody Picture Vocabulary Test (PPVT-III-NL, Dunn and Dunn, 2005). In this test, children were asked to select one out of four picture drawings after an orally presented word. Whereas this task is usually performed as a paper-and-pencil task, stimuli presentation in the current study was controlled by the experimental software E-Prime 2.0 (Schneider et al., 2002), and administered through a laptop computer to facilitate administration and scoring. The shortened version used in our study contained eight items per test set, instead of the usual twelve items, due to testing time constraints. Sets 3, 4, and 5 were presented. As each set contained eight items, there were 24 items in total. Pilot research with 97 three-year-olds established that the items that were removed did not differentiate well among children, as they were either very easy or very difficult (i.e., mean scores on these items were either below 30% or above 70% correct). Scores were calculated as the percentage of correct responses for each child.

Background variables

Socioeconomic status. Parental education was used as an indicator of SES. In two-parent households, parental education of the parent with the highest education was taken as a proxy for family SES. Intermediate vocational education or lower were coded as low to middle SES, while a higher vocational college or University education were coded as high SES. SES information was collected through parent questionnaire at the first study wave; if parent reports were missing at this wave, parents were asked to report SES in subsequent study waves. SES was available for 1843 children (65%).

Home language. Parents reported on children's home language in the parent questionnaire. For the purposes of the present study, we coded whether Dutch was the only language children were exposed to at home or whether they were (also) exposed to (an)other language(s). As the parent questionnaire was missing for a large number of children (see sample description section below), we asked assessors to record children's home language as well at both waves. Assessors were instructed to enquire after children's home language with the parents in the home-based sample and with teachers in the center-based sample. When parent questionnaire data were not available, the wave one assessor's report of home language was imputed. In cases where wave one assessor reports were also unavailable, wave two reports were used. When wave one and wave two assessor reports provided conflicting information, the home language variable was set to missing. Home language information was available for 2463 children (87%).

ANALYTIC STRATEGY

First, we investigated model fit of a one-factor and two-factor (hot vs. cool) EF model using CFA. The fit of the CFA models was assessed with the comparative fit index (CFI) and the root mean square error of approximation (RMSEA) (Kline, 2011). CFI values greater than 0.90 and RMSEA values of less than 0.08 were considered as acceptable fit (Hu and Bentler, 1999). CFI values greater than 0.95 and RMSEA values of less than 0.05 were considered as good fit (Schreiber et al., 2006). As χ^2 is not appropriate for investigating model fit when the sample size is very large, we only report χ^2 for the sake of clarity. The best fitting model was selected for further analyses.

Second, multi-group CFA models were used to evaluate measurement invariance of the EF model, with gender (boys vs. girls), age (below 2.5 years vs. above 2.5 years), home language (monolingual Dutch vs. other), SES (low/middle vs. high), and test setting (home vs. daycare center) as grouping variables. Measurement invariance was investigated by testing the

equivalence of factor loadings and thresholds across groups (Millsap, 2011; Muthén and Muthén, 2013). Four nested models were tested successively for each grouping factor. The first model (configural invariance model) had no constraints regarding any parameter across groups. This model was used to evaluate whether the model held for both groups. In the second model (metric invariance model), factor loadings were constrained to be equal for both groups. For identification purposes, the mean of the reference (first) group was fixed to zero and scale factors were fixed to one. Furthermore, the first threshold value of an indicator was constrained to be equal across groups. The intercept/thresholds of the indicator that was used to set the metric of the model was also constrained to be equal across groups. In the third model (scalar invariance model), all factor loadings, intercept, and thresholds were constrained to be equal across groups. Other settings were equal for the second and third model. In the fourth model (factor covariance model), we constrained the association between the latent factors in the two-factor hot and cool EF model to be equal between groups (Schmitt and Kuljanin, 2008).

As the sample size was large, classical difference testing using the χ^2 was not appropriate (Cheung and Rensvold, 2002; Chen, 2007). Therefore, following recommendations by Chen (2007), we evaluated whether measurement invariance was present by considering changes in CFI and RMSEA. Specifically, a CFI change of 0.01 or less and RMSEA of 0.015 or less indicates measurement invariance for any of the tested sorts; a CFI change above 0.01 and/or RMSEA change exceeding 0.015 indicates measurement invariance is not supported.

Third, we assessed criterion, convergent, and predictive validity of the EF latent factor model in a set of separate analyses. Criterion validity was studied by regressing the latent EF factor(s) on age, gender, home language, SES, and test setting. An alternative approach would have been to compare latent mean factors across groups in the multi-group analyses described above. However, age, home language, SES and test setting were significantly associated with each other. Therefore, a multivariate approach was deemed more appropriate than multigroup comparisons to determine criterion validity. Convergent validity was studied by assessing the association between the latent EF factor(s) and a latent inhibitory control factor, using parent and teacher rated inhibitory control as indicators, and a latent attention factor, using parent attentional focusing and teacher work attitude as indicators. As we were interested in the shared variance within each construct and not in the shared variance within reporters, inhibitory control and attention, reported by parents and teachers, were modeled separately. First, for both the inhibitory control and attention model, separate parent and teacher latent factors were constructed. Next, secondary factors representing the shared variance between parent and teacher reports were modeled and correlated with latent EF. To control for age at assessment, age was entered as a covariate for all latent factors in these models.

Finally, predictive validity was studied by regressing children's latent pre-academic skills, using emergent math skills and vocabulary as indicators, and children's externalizing behavior problems, using parent and teacher ratings as indicators, at age 3 years on the latent EF factor(s) at age 2 years. Age at assessment was controlled for, by regressing the latent EF factor(s) on children's age at wave one, and the latent pre-academic and behavior problem factors on age at wave two. As age at the two waves was significantly associated, the correlation between age at the first and second wave was also included. To make full use of the large dataset, the model was run for the sample as a whole, despite the fact that teacher reports were not available for children in the home cohort. To evaluate whether our findings were robust despite the fact that teacher questionnaire data were missing by design in the full sample, we also tested the model in the center cohort alone.

To investigate the missing data pattern, missingness was analyzed as a function of cohort, home language, and gender. Not enough SES information was available to investigate missingness in relation to SES reliably. We coded missingness on parent and teacher questionnaires and child assessments as the presence or absence of at least one parent questionnaire, teacher questionnaire, and child task score across waves, respectively. Data missingness on child tests and parent questionnaires was significantly associated with cohort and home language, but not gender. There were more missings in children from non-monolingual Dutch families and children in the center cohort on these variables. Missingness on teacher questionnaires was not significantly associated with home language or gender. Given the association between some of the background variables and data missingness, cohort, home language, gender, and SES were entered as covariates in addition to age in each of the validity models. The correlations between these background variables were also included. All available data were used in the analyses; for example, if a child had missing data on one of the EF measures, his or her scores on the other measures were still used in the CFA models. By including the covariates in the validity models, missing data were estimated using the covariates rather than by removing cases, thus preventing estimation bias. All analyses were conducted in Mplus 7.11 (Muthén and Muthén, 1998–2012). As proposed by Byrne and Stewart (2006), WLSMV was used as an estimator in all analyses, because categorical items were present.

RESULTS

SAMPLE DESCRIPTION AND TASK COMPLETION

In total, 2827 children were enrolled in the study. Of those, 390 children (14%) did not complete any of the EF tasks, with task completion defined as responding to at least half of the items of a test. Reasons for not completing a test varied from non-compliance, child illness and language difficulties, to external factors which disturbed the testing situation or technical difficulties. The number of children completing each of the tasks is shown in **Table 1**. Of the 2437 children who completed at least one of the EF tasks, 64% completed all five tasks, 23% completed four tasks, 8% completed three tasks, 2% completed two tasks, and 4% completed one task. At the second wave, vocabulary scores were available for 2088 children and emerging math scores were available for 2063 children (74 and 73% of the full sample, respectively). There were 2604 children (92% of the full sample) for whom at least one task score (EF, emergent math,

Table 1 | Descriptive statistics for executive function measures.

Task	*N*	% completion[a]	*M*	*SD*	Range	Skew (*SE*)	Kurtosis (*SE*)	Floor (%)	Ceiling (%)
Continuous measures									
Visual search	2174	77/89	3.5	1.7	0.3–8	−0.1 (0.1)	−0.8 (0.1)	2.8	0.05
Six boxes	2186	77/90	64.6	18.7	0–100	−0.2 (0.1)	−0.1 (0.1)	0.3	7.6
Memory for location[b]	1803	64/74	2.0	0.9	0–4	0.3 (0.1)	−0.4 (0.1)	2.1	5.4

Categorical measures	*N*	Score distribution (%)			
		% completion[a]	0	1	2
Snack delay	2298	81/94	29.5	21.3	49.3
Gift delay	2289	81/94	17.7	29.5	52.8

[a] Task completion is shown as: percentage of the total sample (N = 2827)/percentage of the sample who completed at least one test (N = 2437). [b] The lower number of children completing the memory for location task was due to the fact that this task was reduced in length after data collection had already begun; data of the first group of children that was assessed were not available for the present analysis.

and/or vocabulary) was available across waves. Parent reports were available for 1471 children at the first wave and 1351 children at the second wave (52 and 48% of the full sample, respectively). There were 1820 children (64%) for whom at least one parent questionnaire was available. Teacher reports were available for 910 children at the first wave and 904 children at the second wave (50% of children in the center sample). There were 1279 children (70% of children in the center sample) for whom at least one teacher questionnaire was available. There were 171 children for whom no data were available on tasks and questionnaires at both measurement waves (6% of the full sample), 129 children for whom task and/or questionnaire data were present at the second, but not at the first wave (5%), 486 children who had task and/or questionnaire data at the first, but not the second wave (18% loss to follow-up), and 2041 children for whom task and/or questionnaire data were present at both waves (72%).

DESCRIPTIVE STATISTICS

Descriptive statistics for each of the EF tasks are shown in **Table 1**. The visual search, six boxes, and memory for location task did not show strong ceiling or floor effects. The categorical delay task measures showed a less optimal distribution, with about half the sample passing each of the tasks (i.e., not touching the snack or gift). At age 3 years, the mean score of the emergent math task was 40.3 ($SD = 10.6$; range = 2.3–72.6) and the mean of the vocabulary task was 63.7 ($SD = 18.2$; range = 0–100).

Table 2 shows the correlations between each of the continuous EF measures. The visual search, six boxes, and memory for location task scores were significantly correlated with each other in the expected direction, although correlations were weak. Each of the measures was also significantly related to age, as expected given the large age range in our sample. When controlling for the effect of age, the correlations between measures were reduced in strength but remained statistically significant. With respect to the categorical EF measures, there was a significant association between the snack and gift delay task scores [$\chi 2_{(4)} = 706.2$; $p < 0.001$]. **Table 3** shows that performance on both the snack and gift delay task was significantly and positively associated to performance on the visual search, six boxes, and memory for location tasks, also after controlling for the shared variance with age.

Table 2 | Correlations between continuous executive function measures.

	Visual search	Six boxes	Memory for location	Age
Visual search	-	0.22***	0.25***	0.36***
Six boxes	0.18***	-	0.17***	0.20***
Memory for location	0.19***	0.13***	-	0.18***

****$p < 0.001$. Correlations below the diagonal are partial correlations corrected for age.*

BASELINE MODEL

Next, we investigated model fit of both a one-factor and two-factor model, with separate cool and hot EF latent factors specified in the latter. In the two-factor model, the visual search, six boxes, and memory for location tasks were indicators of the cool EF factor, while the snack and gift delay tasks were indicators of the hot EF factor. In both models, age at wave one was included as a covariate. The one-factor model showed poor fit [χ^2 (9, $N = 2383$) = 326.58, $p < 0.001$, RMSEA = 0.122 (0.111–0.133), CFI = 0.838]. However, model fit of the two-factor model was good [χ^2 (7, $N = 2383$) = 29.62, $p < 0.001$, RMSEA = 0.037 (0.024–0.051), CFI = 0.988]. For hot EF, standardized factor loadings for the snack delay and gift delay task were 0.77 and 0.86 (p's < 0.001), respectively. For cool EF, standardized factor loadings for the visual search, six boxes, and memory for location task were 0.61, 0.42, and 0.41 (p's < 0.001), respectively. Furthermore, age was a significant predictor of both cool and hot EF ($\beta = 0.53$, $p < 0.001$; $\beta = 0.28$, $p < 0.001$, respectively). Finally, the cool and hot EF factors were significantly associated ($\beta = 0.44$; $p < 0.001$). Because of the better model fit of the one- compared to two-factor model, the two-factor model was used for further analysis.

Table 3 | ANOVA with categorical snack and gift delay task scores as independent variables and continuous executive function measures as dependent variables.

	Snack delay (*M, SD*)			*F* (*df*)	*F* (*df*)[a]
	0	**1**	**2**		
Visual search	3.2 (1.6)	3.4 (1.6)	3.8 (1.6)	26.7 (2, 2053)***	10.4 (2, 2050)***
Six boxes	60.5 (19.8)	62.7 (19.1)	67.7 (17.4)	32.8 (2, 2127)***	21.6 (2, 2079)***
Memory for location	1.8 (0.9)	2.0 (0.9)	2.1 (0.9)	20.9 (2, 1746)***	11.9 (2, 1732)***
	Gift delay (*M, SD*)				
	0	**1**	**2**		
Visual search	2.8 (1.7)	3.2 (1.7)	3.7 (1.7)	47.7 (2, 2059)***	25.9 (2, 2056)***
Six boxes	59.5 (19.7)	62.0 (18.6)	67.6 (17.9)	34.9 (2, 2144)***	24.3 (2, 2093)***
Memory for location	1.7 (0.9)	1.9 (0.9)	2.1 (0.9)	23.5 (2, 1746)***	15.5 (2, 1730)***

[a]*ANCOVA analysis with age as covariate.* ***$p < 0.001$.

MEASUREMENT INVARIANCE

Next, we investigated whether the two-factor hot and cool EF model showed measurement invariance across subgroups of age, gender, home language, SES, and test setting. For each of these grouping variables, a set of nested models was tested and compared to each other, after constraining an increasing number of parameters. Age was controlled for in all models, except in the model where age was the grouping variable. Model fit was good for all models (CFI > 0.95, RMSEA < 0.05; **Table 4**). Configural, metric, scalar and factor covariance invariance was supported across all subgroups, as the changes in CFI were never larger than 0.01 and the changes in RMSEA never exceeded 0.015. Thus, the two-factor hot and cool EF model fitted the data well in all groups, and factor loadings and intercepts (continuous variables) and thresholds (categorical variables) of the indicators could be constrained to equality between groups differing in age, gender, home language, SES, and test setting. In addition, the association between the hot and cool factors could be constrained to equality between groups. In sum, the two-factor hot and cool EF model showed strong measurement invariance across age, gender, home language, SES, and test setting groups.

CRITERION VALIDITY

To investigate criterion validity, latent hot and cool EF factors were regressed on age, gender, SES, home language, and test setting. Model fit was good [χ^2 (19, $n = 2827$) = 57.23, $p < 0.001$, RMSEA = 0.027 (0.019–0.035), CFI = 0.987]. Age was positively related to both cool and hot EF, so that older children obtained higher scores than younger children ($\beta = 0.61$, $p < 0.001$; $\beta = 0.25$, $p < 0.001$, respectively). Also, girls obtained higher scores than boys on both cool and hot EF ($\beta = 0.16$, $p < 0.001$; $\beta = 0.10$, $p < 0.001$, respectively). Although SES was positively related to cool EF, no effect of SES on hot EF was observed ($\beta = 0.23$, $p < 0.001$; $\beta = 0.03$, $p = 0.313$, respectively). Children from monolingual Dutch families obtained higher cool and hot EF scores than children from families in which another language next to or instead of Dutch was spoken ($\beta = 0.19$, $p < 0.001$; $\beta = 0.08$, $p = 0.004$, respectively). Furthermore, children who were tested at their daycare center had higher scores on hot EF than children who were tested at home ($\beta = 0.12$, $p < 0.001$). No effect of test setting on cool EF was observed ($\beta = -0.009$, $p = 0.783$).

CONVERGENT VALIDITY

To evaluate convergent validity of the test battery, the associations between the latent hot and cool EF factors and parent and teacher reports of children's inhibitory control and attention were studied. The model validating the EF assessment against report-based inhibitory control had acceptable fit [χ^2 (127, $N = 2827$) = 348.02, $p < 0.001$, CFI = 0.946, RMSEA = 0.025 (0.022–0.028)]. Both hot and cool EF were significantly and positively related to report-based inhibitory control (see **Figure 1A**). The association between hot EF and report-based inhibitory control was not significantly different from the association between cool EF and report-based inhibitory control [ω (1) = 0.29, $p = 0.588$]. The model validating the EF assessment against report-based attention fitted the data well [χ^2 (110, $N = 2827$) = 195.26, $p < 0.001$, CFI = 0.978, RMSEA = 0.017 (0.013–0.020)]. Both hot and cool EF latent factors were significantly positively related to report-based attention (see **Figure 1B**). However, the association with report-based attention was larger for cool compared to hot EF [ω (1) = 11.02, $p < 0.001$]. When only children in the center sample were included in the analyses, both models fitted the data well [report-based inhibitory control model: χ^2 (116, $N = 1802$) = 219.00, $p < 0.001$, CFI = 0.949, RMSEA = 0.022 (0.018–0.027); report-based attention model: χ^2 (100, $N = 1802$) = 159.73, $p < 0.001$, CFI = 0.973, RMSEA = 0.018 (0.013–0.023)]. The same pattern of results was found [report-based inhibitory control and hot EF: $\beta = 0.33$, $p = 0.003$; cool EF: $\beta = 0.28$, $p = 0.034$; ω (1) = 0.11, $p = 0.741$; report-based attention and hot EF: $\beta = 0.22$, $p = 0.016$; cool EF: $\beta = 0.59$, $p < 0.001$; ω (1) = 7.00, $p = 0.008$].

PREDICTIVE VALIDITY

In a final set of analyses, we investigated the predictive validity of the hot and cool EF constructs at age 2 years for behavioral functioning and pre-academic skills at age 3 years.

Table 4 | Analysis of measurement invariance across groups.

Model		Model fit indices			Nested model comparisons		
		χ2 (*df*)	CFI	RMSEA	Comp	ΔCFI	ΔRMSEA
AGE (*N* = 2383)							
1a	Configural	16.62* (8)	0.994	0.030			
1b	Metric	13.89 (11)	0.998	0.015	vs. 1a	0.004	0.015
1c	Scalar	29.06* (14)	0.990	0.030	vs. 1b	0.008	0.015
1d	Factor covariance	32.40** (15)	0.989	0.031	vs. 1c	0.001	0.001
GENDER (*N* = 2425)							
2a	Configural	31.70** (14)	0.991	0.032			
2b	Metric	27.73* (17)	0.994	0.023	vs. 2a	0.003	0.009
2c	Scalar	30.17 (20)	0.995	0.020	vs. 2b	0.001	0.003
2d	Factor covariance	29.28 (21)	0.996	0.018	vs. 2c	0.001	0.002
HOME LANGUAGE (*N* = 2350)							
3a	Configural	31.88** (14)	0.991	0.033			
3b	Metric	41.93*** (17)	0.987	0.035	vs. 3a	0.004	0.002
3c	Scalar	47.28*** (20)	0.986	0.034	vs. 3b	0.001	0.001
3d	Factor covariance	49.49*** (21)	0.985	0.034	vs. 3c	0.001	<0.001
SES (*N* = 1772)							
4a	Configural	29.03* (14)	0.990	0.035			
4b	Metric	26.06 (17)	0.994	0.025	vs. 4a	0.004	0.010
4c	Scalar	30.92 (20)	0.992	0.025	vs. 4b	0.002	<0.001
4d	Factor covariance	32.28 (21)	0.992	0.025	vs. 4c	<0.001	<0.001
TEST SETTING (*N* = 2524)							
5a	Configural	28.34*(14)	0.993	0.028			
5b	Metric	39.53** (17)	0.988	0.032	vs. 5a	0.005	0.004
5c	Scalar	62.14*** (20)	0.978	0.041	vs. 5b	0.010	0.009
5d	Factor covariance	72.64*** (21)	0.973	0.044	vs. 5c	0.005	0.003

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Parent and teacher ratings on the BITSEA externalizing behavioral problem scale items were used as indicators to a latent multi-informant preschool externalizing behavior problem factor. Children's vocabulary and emerging math skills test scores were used to create a latent preschool pre-academic score. The predictive validity model had acceptable fit [χ^2 (193, N = 2827) = 555.37, $p < 0.001$, CFI = 0.950, RMSEA = 0.026 (0.023–0.028)], as shown in **Figure 2**. Wave one cool EF was a significant predictor of both preschool externalizing behavior problems and emergent math and vocabulary as indicators of children's pre-academic skills at wave two. In contrast, wave one hot EF was a significant predictor of externalizing behavior problems, but not pre-academic skills, at wave two. The observed effects were unique effects: the effects of cool EF held while controlling for hot EF, and vice versa. When only children in the center cohort were included, model fit was acceptable χ^2 (180, $N = 1798$) = 389.31, $p < 0.001$, CFI = 0.954, RMSEA = 0.025 (0.022–0.029). Results were similar to those of the full sample; the only difference was that the effect of hot EF on externalizing behavior problems was now no longer significant [problem behavior on hot EF: $\beta = -0.15$, $p = 0.081$; cool EF: $\beta = -0.19$, $p = 0.038$; pre-academic skills on hot EF: $\beta = 0.01$, $p = 0.857$; cool EF: $\beta = 0.32$, $p < 0.001$).

DISCUSSION

Executive function is an important predictor of academic achievement (Blair and Diamond, 2008), socio-emotional development (Carlson et al., 2004), and behavioral adjustment (Eisenberg et al., 2009; Espy et al., 2011) in the preschool period and beyond. The lack of psychometrically well-validated EF assessment instruments for very young children is a major obstacle to further progress our understanding of EF development in the early years (Blair and Ursache, 2010). The present study aimed to fill this void by investigating the psychometric quality of an executive function test battery for two-year-olds using confirmatory factor analysis. The EF task battery used in this study included measures of both cool EF (working memory, attention) and hot EF (delay of gratification). The battery comprised both new measures which were developed for the purposes of this study, and existing measures which were adapted for use in a large-scale field study. CFA analyses showed that (1) a two-factor hot and cool EF model fitted the data better than a one-factor EF model, (2) measurement invariance was supported across different subgroups of age, gender, home language, SES, and test setting, and (3) the test battery showed satisfactory criterion, convergent, and predictive validity.

Our first finding that a two-factor hot and cool EF model fitted the data better than a one-factor model is in line with results

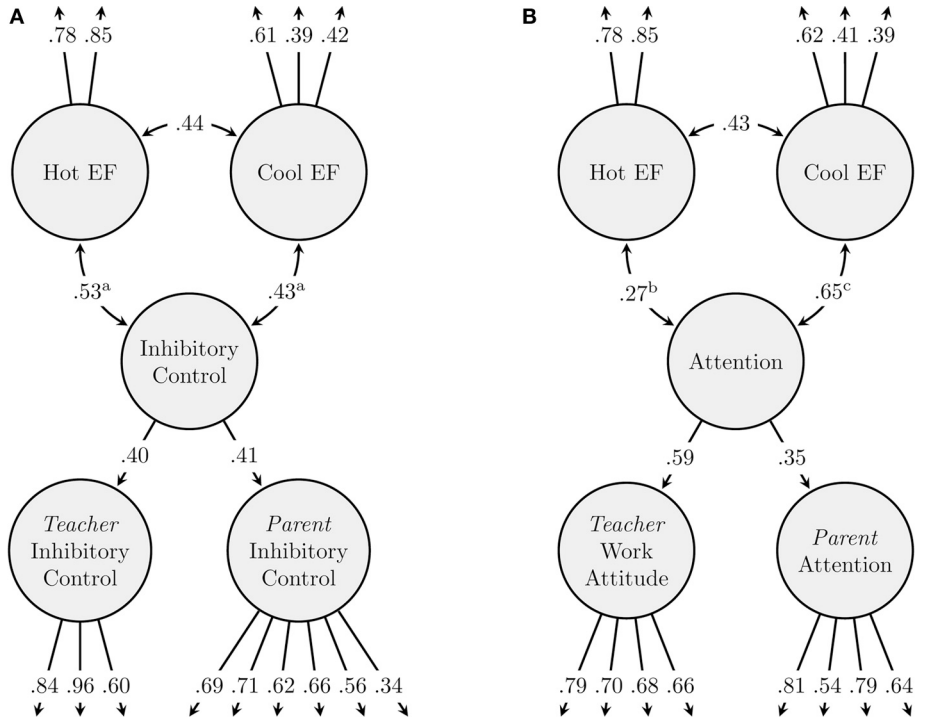


FIGURE 1 | (A) Convergent validity model: Association between cool and hot EF and report-based inhibitory control. Age, gender, SES, home language, and cohort were included as covariates in the model. **(B)** Convergent validity model: Association between cool and hot EF and report-based attention. Age, gender, SES, home language, and cohort were included as covariates in the model.

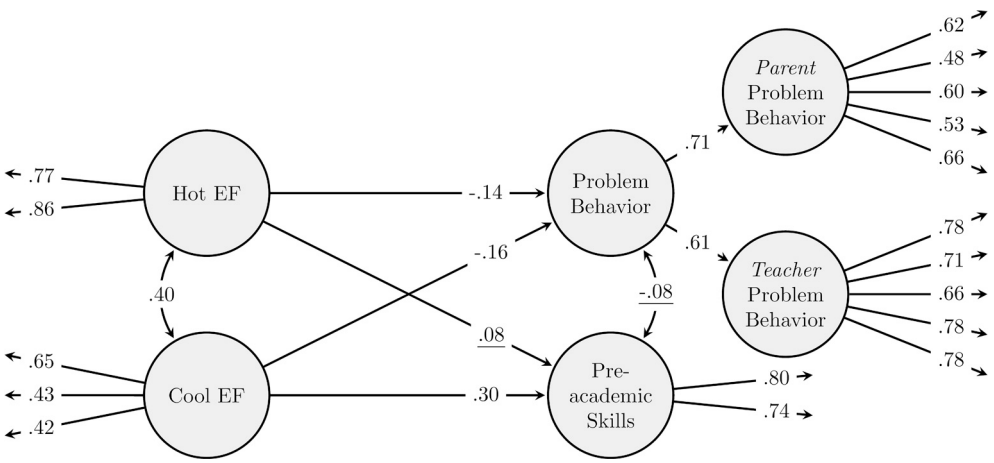


FIGURE 2 | Predictive validity model. Age, gender, SES, home language, and cohort were included as covariates in the model. Values which are underlined are not significant.

of most previous studies on preschoolers (Brock et al., 2009; Willoughby et al., 2011; Bassett et al., 2012; Kim et al., 2013). To the best of our knowledge, the current study is the first to provide support for a distinction between hot and cool EF in children as young as age 2 years through using CFA. However, it should be noted that the two indicators of the hot EF latent factor, the gift and snack delay task, were very similar in instruction and format. As such, these tasks may have loaded strongly on the same latent factor due to other factors than children's executive processing of affective information only. Future research could explore if similar results are obtained if more differentiated measures of hot EF are used with very young children.

Additional evidence for the differentiation between hot and cool EF factors at this young age comes from our predictive validity analyses. These analyses showed that the latent cool and hot EF factors were differentially predictive of children's outcomes 1

year later. In particular, the cool EF latent factor predicted both emergent math and vocabulary and externalizing behavior problems at age 3 years, whereas the hot EF factor only predicted externalizing behavior problems. These results support previous research showing similar relationships (Willoughby et al., 2011; Bassett et al., 2012; Kim et al., 2013; but see Brock et al., 2009). Note however, that our results were somewhat mixed: the association between hot EF and externalizing behavior problems was only observed in the full sample, and not in the center sample alone. In both analyses, the effect was relatively weak in strength. This could perhaps be explained by the fact that the two delay of gratification measures, which were used as indicators to the hot EF constructs, were not optimally distributed. In particular, in both tasks, about half of the children obtained the highest score. As such, there was not much differentiation between children at the higher end of the ability spectrum for hot EF, which could perhaps explain the relatively weak association with outcome measures.

Besides a CFA analysis comparing a one- and two-factor model, we tested measurement invariance of the latter, preferred, model across a number of different subgroups: younger vs. older toddlers, boys vs. girls, monolingual Dutch vs. other language groups, low/middle vs. high SES, and home-based assessment vs. center-based assessment settings. Strong measurement invariance was found, as the factor structure, factor loadings, intercepts of indicators, and associations between the hot and cool latent factors could all be constrained to equality across groups. This indicates that the tasks in our battery tap underlying EF ability in the same way in different subgroups of children, thus allowing for a fair comparison of latent hot and cool EF ability across these subgroups.

A further set of analyses showed significant relations between cool and hot EF, on the one hand, and gender, SES, age, home language, and test setting, on the other, supporting the criterion validity of the test battery. With respect to gender, girls outperformed boys on both EF constructs. Cross-cultural comparative studies have shown that child gender differences in EF, using the Head-Toes-Knees-Shoulders task (Ponitz et al., 2009) occur in some countries (i.e., the United States and Iceland), but not others (i.e., Taiwan, China, South Korea, Germany, and France) (Wanless et al., 2013; Gestsdottir et al., 2014). In the Netherlands, Huizinga and Smidts (2011) found higher EF scores for girls compared to boys in Dutch school-aged children and adolescents, using parent reports of EF. The results of the present study add to these findings by showing that such gender differences are already present well before school age, using a different assessment method, i.e., direct child assessments. Furthermore, we observed an effect of test setting on hot, but not cool EF. Children assessed at home achieved lower hot EF scores than children tested at their daycare center, after controlling for age, gender, SES, and home language. It is unclear, however, which factors could explain this effect. We also observed that children from non-monolingual Dutch homes scored lower on both cool and hot EF than their monolingual Dutch peers. The main subgroup in the non-monolingual Dutch sample consisted of non-Western immigrants, but results from the current study in this domain should be interpreted with caution, as this sample was very mixed. Our findings may indicate that differences in child rearing practices in different cultural groups can impact on EF development already at this young age. A recent cross cultural comparison across Syrian and German five- to twelve-year-old children showed that Syrian children performed less well on measures of sustained attention, visuospatial orienting, and executive function than their German peers (Sobeh and Spijkers, 2013). Alternatively, our findings could be due to differences in linguistic abilities across groups. Future studies are needed to investigate how cultural differences and associated child rearing practices, as well as linguistic differences, impact on EF development.

Apart from effects of age, gender and home language, an effect of SES was found such that children from low/middle SES backgrounds scored lower on cool EF than their high SES peers. The negative impact of low SES on EF in older children is well established (Noble et al., 2005, 2007). The present results add to these findings and show an effect of SES at a younger age, corroborating the findings by Hughes and Ensor (2005) that SES is related to EF at toddler age already. As previous studies have shown that a gap in academic achievement between children from low and high SES families persists over time (Heckman, 2006; Hackman et al., 2010), it seems especially important to design interventions to promote EF development in low SES children at a very young age. To date, preschool and parenting programs have most often focused on promoting EF in preschoolers, i.e., three- to five-year-olds, from disadvantaged families (e.g., Diamond et al., 2007; Neville et al., 2013). However, a recent review showed that attentional control and working memory training leads to more widespread transfer effects when given to younger children, potentially due to the fact that neural plasticity is larger in younger children (Wass et al., 2012). As such, there is a need to develop effective interventions to promote EF development in disadvantaged children even before preschool age and to design curricula for center-based education and care for young children that foster EF development.

In contrast to our findings regarding the influence of SES on cool EF, no SES effect was apparent on hot EF as measured with the snack and gift delay tasks. Previous studies have reported conflicting results regarding the role of SES in performance on delay tasks. For example, Li-Grining (2007) showed that there was no effect of socio-demographic risk on preschoolers' delay of gratification. In contrast, Evans and English (2002) found that eight- to ten-year-olds from low-income families performed less well on a delay of gratification measure than their peers from middle-income families. Thus, more research is necessary to investigate the role of SES on the development of delay of gratification, and whether effects of SES are specific to certain types of delay tasks or age ranges.

Finally, we found moderate correlations between the hot and cool EF factors and parent and teacher reports of children's attention and inhibitory control (Rothbart et al., 2003), supporting the EF tasks' convergent validity. Divergent validity was supported by the stronger association between reports of attention and children's cool EF compared to hot EF ability. Parent and teacher reports of children's attention mostly included items which covered the ability to remain focused and concentrate for longer

periods of time. It is clear that such attentional focusing behavior was an important prerequisite for performing well in the working memory and selective attention measures. However, previous studies have shown that attention deployment is also an important factor in delay of gratification, or hot EF. Although we found evidence for this relation too, the association between hot EF and reported attention was weaker than between cool EF and reported attention. Potentially, in addition to the ability to remain focused, an alternative mechanism of selective attentional deactivation or distraction is more important for hot EF. For example, in the classic "marshmallow test," children who distract themselves effectively from the single marshmallow which is put in front of them, are more effective at delaying gratification and waiting for a larger reward (i.e., two marshmallows at a later time), than children who focus on the single marshmallow instead (Mischel and Ebbesen, 1970; Peake et al., 2002). Future research is needed to investigate the association between these two types of attention deployment in different situations. The latent cool and hot EF factors were equally strongly related to reports of inhibitory control. This finding is not surprising, given that all three cool EF measures required some form of inhibitory control as well. In the selective attention task, children had to suppress pointing to distracting animals. Also, the six boxes visuospatial working memory task (Diamond et al., 1997) and memory for location task required children to search for hidden toys in identical boxes, and not re-open the boxes they had just opened. We observed that some toddlers sometimes made perseveration errors on these tasks, suggesting that inhibitory control processes play a role in performance, as has previously been observed in other search tasks for young children, such as the A-not-B task (Diamond et al., 1994).

The present study contributes to the extant literature about EF in early childhood in a number of ways. It is, to the best of our knowledge, one of the few validation studies to date that focused on children as young as 2 years, supporting the validity of an EF assessment already at this young age. Moreover, it used CFA to investigate whether the current EF assessments represented a one- or two-factor structure, and thoroughly investigated measurement invariance across various subgroups. Importantly, unlike previous studies, our study was conducted in a large, nationwide sample, involving over 2000 children and including a large number of children from low/middle SES families and families in which other languages than Dutch were spoken, increasing the external validity of the results. Furthermore, children's EF task measures were triangulated by independent parent and teacher reports on children's behavior in naturalistic settings at home and in daycare, revealing considerable shared variance, supporting the validity of the EF measures. Moreover, we assessed predictive validity of the test battery, showing significant associations between children's EF at age 2 years and behavioral problems and pre-academic skills (i.e., vocabulary and emergent math) at age 3 years.

There are, however, also a number of limitations. First of all, missing data were substantial, especially regarding SES and parent and teacher reports of children's behavior. Second, it would have been beneficial to include more tasks in each domain. In the cool EF domain, inclusion of measures of shifting and inhibitory control would have allowed for a more comprehensive construct. However, in our experience, selecting tasks to assess shifting and inhibitory control for such young children is challenging. These tasks often rely on "if-then" rules (e.g., Go-NoGo tasks in which the child is instructed to press a key if stimulus X is shown and withhold their response when stimulus Y is shown) and such rules are often too challenging for two-year-olds (Zelazo and Reznick, 1991), although shifting measures have been successfully administered to two-year-olds in some studies (e.g., Hughes and Ensor, 2007; Beck et al., 2011). In the hot EF domain, more delay tasks with a different administration format would have provided a more pure hot EF latent construct. However, decisions regarding measurement selection were made with testing time in mind; for the purposes of this large-scale field study with very young children, test time was limited. Finally, we used non-standard versions of the ECBQ and BITSEA questionnaires for validation purposes.

A number of implications arise from the current study. First, our results showed that the assessment of EF through using multiple measures and modeling latent constructs showed satisfactory to good psychometric properties in this very young sample. Thus, the current study shows that EF can be assessed reliably in children as young as 2 years of age. As in the study by Willoughby et al. (2010) with three-year-olds, we observed that the associations among separate EF tasks were relatively weak. However, all tasks loaded significantly onto their latent factors, and latent factors were in turn significantly related to a number of outcome measures, with substantial effect sizes. Our findings thus support Willoughby et al.'s (2010) conclusion that, especially for very young children, it is recommended to use multiple EF measures to be able to construct latent factors. This way, the influence of measurement error is reduced and the reliability of the EF assessment is increased (Willoughby and Blair, 2011). Second, the current study shows that, even at the age of 2 years, EF can be meaningfully differentiated in a cool and hot component. Thus, for applied research in which an assessment of EF is included to predict children's outcomes across multiple domains, inclusion of both hot and cool EF measures is recommended.

To conclude, our EF task battery for two-year-old children in the Netherlands showed satisfactory psychometric quality and criterion, convergent and predictive validity. We are currently investigating data from the children in this sample at age three, four, and five years, to investigate their development of EF from the toddler through to the preschool years. The current instruments offer new opportunities for investigating EF development in early childhood and for evaluating interventions targeted at improving EF from a young age.

AUTHOR CONTRIBUTIONS

Hanna Mulder developed, piloted, and implemented the test battery; wrote introduction, sample, and task descriptions, and discussion; reviewed and revised analyses and results section. Huub Hoofs provided an initial draft of the manuscript as his MSc thesis; wrote analyses and results section. Josje Verhagen co-developed, piloted and implemented the test battery; reviewed and revised the manuscript. Ineke van der Veen project design; general project methodology; questionnaire design. Paul P. M.

Leseman principal investigator of the project; reviewed and revised the manuscript.

ACKNOWLEDGMENTS

The Pre-COOL study was conducted in collaboration between the Department of Special Education at Utrecht University, the Kohnstamm Institute at the University of Amsterdam, and the Institute for Applied Social Sciences (ITS) at the Radboud University Nijmegen. We are grateful to our project partners at the ITS. The Pre-COOL study is funded by the Dutch research council NWO (grant number 411-20-442). The authors thank Jos Keuning (CITO) for scaling the emergent math scores using item response theory. We are also grateful to all the children, families and day care centers who participated in our study.

REFERENCES

Allan, N. P., and Lonigan, C. J. (2011). Examining the dimensionality of effortful control in preschool children and its relation to academic and socioemotional indicators. *Dev. Psychol.* 47, 905–915. doi: 10.1037/a0023748

Bassett, H. H., Denham, S., Wyatt, T. M., and Warren-Khot, H. K. (2012). Refining the preschool self-regulation assessment for use in preschool classrooms. *Infant Child Dev.* 21, 596–616. doi: 10.1002/icd.1763

Beck, D. M., Schaefer, C., Pang, K., and Carlson, S. M. (2011). Exeutive function in preschool children: test-retest reliability. *J. Cogn. Dev.* 12, 169–193. doi: 10.1080/15248372.2011.563485

Blair, C., and Diamond, A. (2008). Biological processes in prevention and intervention: the promotion of self-regulation as a means of preventing school failure. *Dev. Psychopathol.* 20, 899–911. doi: 10.1017/S0954579408000436

Blair, C., Granger, D., and Peters Razza, R. (2005). Cortisol reactivity is positively related to executive function in preschool children attending Head Start. *Child Dev.* 76, 554–567. doi: 10.1111/j.1467-8624.2005.00863.x

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Blair, C., and Ursache, A. (2010). "A bidirectional model of executive functions and self-regulation," in *Handbook of Self-Regulation: Research, Theory, and Applications 2nd Edn*, eds K. D. Vohs and R. F. Baumeister, (New York, NY: Guilford), 300–320.

Bleichrodt, N., Resing, W. C. M., and Zaal, J. N. (1993). *SCHOBL-R: Beoordeling van Schoolgedrag.* Amsterdam: Harcourt Test Publishers.

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., and Leseman, P. P. M. (in press). The benefits of being bilingual: working memory in bilingual Turkish-Dutch children. *J. Exp. Child Psychol.*

Briggs-Gowan, M. J., and Carter, A. S. (2001). *BRIEF-Infant Toddler Social and Emotional Assessment (BITSEA) Manual, Version 1.0.* New Haven, CT: Yale University.

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., and Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Child. Res. Q.* 24, 337–349. doi: 10.1016/j.ecresq.2009.06.001

Byrne, B. M., and Stewart, S. M. (2006). Teacher's corner The MACS approach to testing for multigroup invariance of a second-order structure: a walk through the process. *Struct. Equ. Model.* 13, 287–321. doi: 10.1207/s15328007sem1302_7

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M., Mandell, D. J., and Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Dev. Psychol.* 40, 1105–1122. doi: 10.1037/0012-1649.40.6.1105

Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298. doi: 10.1111/j.1467-7687.2008.00675.x

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Modeling* 14, 464–504. doi: 10.1080/10705510701301834

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Modeling* 9, 233–255. doi: 10.1207/S15328007SEM0902_5

Clark, C. A., Pritchard, V. E., and Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Dev. Psychol.* 46, 1176–1191. doi: 10.1037/a0019672

Davidson, M. C., Amso, D., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006

Denham, S. A., Warren-Khot, H. K., Bassett, H. H., Wyatt, T., and Perna, A. (2012). Factor structure of self-regulation in preschoolers: testing models of a field-based assessment for predicting early school readiness. *J. Exp. Child Psychol.* 11, 386–404. doi: 10.1016/j.jecp.2011.10.002

Diamond, A., Barnett, W. S., Thomas, J., and Munro, S. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Diamond, A., Cruttenden, L., and Neiderman, D. (1994). AB with multiple wells: I. Why are multiple wells sometimes easier than two wells? II. Memory or memory + inhibition? *Dev. Psychol.* 30, 192–205. doi: 10.1037/0012-1649.30.2.192

Diamond, A., Prevor, M. B., Callender, G., and Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monogr. Soc. Res. Child Dev.* 62, i-206. doi: 10.2307/1166208

Driessen, G., Mulder, L., Ledoux, G., Roeleveld, J., and Van Der Veen, I. (2009). *Cohortonderzoek COOL5-18, Technisch rapport basisonderwijs, eerste meting 2007/08.* Nijmegen: ITS/SCO-Kohnstamm Institute.

Dunn, M., and Dunn, L. M. (2005). *Peabody Picture Vocabulary Test-III-NL.* Amsterdam: Harcourt Test Publishers

Eisenberg, N., Valiente, C., Spinrad, T. L., Liew, J., Zhou, Q., Losoya, S. H., et al. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to their externalizing, internalizing, and co-occurring behavior problems. *Dev. Psychol.* 45, 988–1008. doi: 10.1037/a0016213

Espy, K. A., Sheffield, T. D., Wiebe, S. A., Clark, C. A., and Moehr, M. J. (2011). Executive control and dimensions of problem behaviors in preschool children. *J. Child Psychol. Psychiatry* 52, 33–46. doi: 10.1111/j.1469-7610.2010.02265.x

Evans, G. W., and English, K. (2002). The environment of poverty: multiple stressor exposure, psychophysiological stress, and socioemotional adjustment. *Child Dev.* 73, 1238–1248. doi: 10.1111/1467-8624.00469

Fitzpatrick, C., and Pagani, L. S. (2012). Toddler working memory skills predict kindergarten school readiness. *Intelligence* 40, 205–212. doi: 10.1016/j.intell.2011.11.007

Garon, N., Bryson, S. E., and Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi: 10.1037/0033-2909.134.1.31

Garon, N., Smith, I. M., and Bryson, S. E. (2013). A novel executive function battery for preschoolers: Sensitivity to age differences. *Child Neuropsychol.* doi: 10.1080/09297049.2013.857650. [Epub ahead of print].

Gerhardstein, P., and Rovee-Collier, C. (2002). The development of visual search in infants and very young children. *J. Exp. Child Psychol.* 81, 194–215. doi: 10.1006/jecp.2001.2649

Gestsdottir, S., von Schudoletz, A., Wanless, S. B., Hubert, B., Guimard, P., Birgisdottir, F., et al. (2014). Early behavioral self-regulation, academic achievement, and gender: Longitudinal findings from France, Germany, and Iceland. *Appl. Dev. Sci.* 18, 90–109. doi: 10.1080/10888691.2014.894870

Hackman, D. A., Farah, M. J., and Meaney, M. J. (2010). Science and Society. Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nat. Rev. Neurosci.* 11, 651–659. doi: 10.1038/nrn2897

Heckman, J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science* 312, 1900–1902. doi: 10.1126/science.1128898

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6, 1–55. doi: 10.1080/10705519909540118

Hughes, C., and Ensor, R. (2005). Executive function and theory of mind in 2 year olds: a family affair? *Dev. Neuropsychol.* 28, 645–668. doi: 10.1207/s15326942dn2802_5

Hughes, C., and Ensor, R. (2007). Executive function and theory of mind: predictive relations from ages 2 to 4. *Dev. Psychol.* 43, 1447–1459. doi: 10.1037/0012-1649.43.6.1447

Huizinga, M., and Smidts, D. P. (2011). Age-related changes in executive function: a normative study with the Dutch version of the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychol.* 17, 51–66. doi: 10.1080/09297049.2010.509715

Kim, S., Nordling, J. K., Yoon, J. E., Boldt, L. J., and Kochanska, G. (2013). Effortful control in "hot" and "cool" tasks differentially predicts childrens behavior problems and academic performance. *J. Abnorm. Child Psychol.* 41, 43–56. doi: 10.1007/s10802-012-9661-4

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford Press.

Kochanska, G., and Knaack, A. (2003). Effortful control as a personality characteristic of young children: antecedents, correlates, and consequences. *J. Pers.* 71, 1087–1112. doi: 10.1111/1467-6494.7106008

Kochanska, G., Murray, K. T., and Harlan, E. T. (2000). Effortful control in early childhood: continuity and change, antecedents, and implications for social development. *Dev. Psychol.* 36, 220–232. doi: 10.1037/0012-1649.36.2.220

Lehto, J. E., Juujärvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive function: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Li-Grining, C. P. (2007). Effortful control among low-income preschoolers in three cities: stability, change, and individual differences. *Dev. Psychol.* 43, 208–221. doi: 10.1037/0012-1649.43.1.208

Mcclelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., and Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Dev. Psychol.* 43, 947–959. doi: 10.1037/0012-1649.43.4.947

Meltzoff, A., Gopnik, A., and Repacholi, B. (1999). "Toddlers' understanding of intentions, desires and emotions: explorations of the dark ages," in *Developing Theories of Intention: Social Understanding and Self-Control*, eds P. Zelazo, J. Astington, and D. Olson (Mahwah, NJ: Lawrence Erlbaum Associates), 17–41.

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., and Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *J. Cogn. Dev.* 13, 395–423. doi: 10.1080/15248372.2011.585478

Miller, S. E., and Marcovitch, S. (2011). Toddlers benefit from labeling on an executive function search task. *J. Exp. Child Psychol.* 108, 580–592. doi: 10.1016/j.jecp.2010.10.008

Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. New York, NY: Routledge.

Mischel, W., and Ebbesen, E. (1970). Attention in delay of gratification. *J. Pers. Soc. Psychol.* 16, 329–337. doi: 10.1037/h0029815

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Muthén, L. K., and Muthén, B. O. (1998–2012). *Mplus User's Guide*, *7th Edn*. Los Angeles, CA: Muthén and Muthén.

Muthén, L. K., and Muthén, B. O. (2013). *Version 7.1. Mplus Language Addendum*. Available online at: https://www.statmodel.com/download/Version7.1x Language.pdf

Neville, H. J., Stevens, C., Pakulak, E., Bell, T. A., Fanning, J., Klein, S., et al. (2013). Family-based training program improves brain function, cognition, and behavior in lower socioeconomic status preschoolers. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12138–12143. doi: 10.1073/pnas.1304437110

Noble, K. G., Mccandliss, B. D., and Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Dev. Sci.* 10, 464–480. doi: 10.1111/j.1467-7687.2007.00600.x

Noble, K. G., Norman, M. F., and Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Dev. Sci.* 8, 74–87. doi: 10.1111/j.1467-7687.2005.00394.x

Op den Kamp, M., and Keuning, J. (2011). *Wetenschappelijke Verantwoording Van de Toets Rekenen Voor Peuters*. Arnhem: Cito.

Peake, P., Hebl, M., and Mischel, W. (2002). Strategic attention deployment for delay of gratification in working and waiting situations. *Dev. Psychol.* 38, 313–326. doi: 10.1037/0012-1649.38.2.313

Pelphrey, K. A., Reznick, J. S., Davis Goldman, B., Sasson, N., Morrow, J., Donahoe, A., et al. (2004). Development of visuospatial short-term memory in the second half of the 1st year. *Dev. Psychol.* 40, 836–851. doi: 10.1037/0012-1649.40.5.836

Ponitz, C. C., McClelland, M. M., Matthews, J. S., and Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Dev. Psychol.* 45, 605–619. doi: 10.1037/a0015365

Poulin-Dubois, D., Blaye, A., Coutya, J., and Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *J. Exp. Child Psychol.* 108, 567–579. doi: 10.1016/j.jecp.2010.10.009

Putnam, S. P., Gartstein, M. A., and Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: the early childhood behavior questionnaire. *Infant Behav. Dev.* 29, 386–401. doi: 10.1016/j.infbeh.2006.01.004

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., and Pressler, E. (2011). CSRP's Impacton low-income preschoolers' preacademic skills: self-regulation as a mediating mechanism. *Child Dev.* 82, 362–378. doi: 10.1111/j.1467-8624.2010.01561.x

Rose, S. A., Feldman, J. F., and Jankowski, J. J. (2009). A cognitive approach to the development of early language. *Child Dev.* 80, 134–150. doi: 10.1111/j.1467-8624.2008.01250.x

Rothbart, M. K., Ellis, L. K., Rosario Rueda, M., and Posner, M. I. (2003). Developing mechanisms of temperamental effortful control. *J. Pers.* 71, 1113–1144. doi: 10.1111/1467-6494.7106009

Scerif, G., Cornish, K., Wilding, J., Driver, J., and Karmiloff-Smith, A. (2004). Visual search in typically developing toddlers and toddlers with Fragile X or Williams syndrome. *Dev. Sci.* 7, 116–130. doi: 10.1111/j.1467-7687.2004.00327.x

Schmitt, N., and Kuljanin, G. (2008). Measurement invariance: review of practice and implications. *Human Resour. Manag. Rev.* 18, 210–222. doi: 10.1016/j.hrmr.2008.03.003

Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., and King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: a review. *J. Educ. Res.* 99, 323–338. doi: 10.3200/JOER.99.6.323-338

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S.-C., and Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Dev. Neuropsychol.* 35, 679–697. doi: 10.1080/87565641.2010.508546

Sobeh, J., and Spijkers, W. (2013). Development of neuropsychological functions of attention in two cultures: a cross-cultural study of attentional performances of Syrian and German children of pre-school and school age. *Eur. J. Dev. Psychol.* 10, 318–336. doi: 10.1080/17405629.2012.674761

Usai, M. C., Viterbori, P., Traverso, L., and de Franchis, V. (2014). Latent structure of executive function in five- and six-year-old children: a longitudinal study. *Eur. J. Dev. Psychol.* 11, 447–462. doi: 10.1080/17405629.2013.840578

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., and Leseman, P. P. M. (2012). The development of executive functions and early mathematics: a dynamic relationship. *Br. J. Educ. Psychol.* 82, 100–119. doi: 10.1111/j.2044-8279.2011.02035.x

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., and Leseman, P. P. M. (2013). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *Br. J. Dev. Psyhol.* 31, 70–87. doi: 10.1111/j.2044-835X.2012.02079.x

Veen, A., Veen, I., Heurter, A., Ledoux, G., Mulder, L., Paas, T., et al. (2012). *Pre-COOL Cohortonderzoek. Technisch Rapport Tweejarigencohort, Eerste Meting 2010-2011*. [Pre-COOL cohort study. Technical report two-year-olds cohort, first measurement 2010-2011]. Amsterdam: Kohnstamm Instituut.

Vicari, S., Caravale, B., Carlesimo, G. A., Casadei, A. M., and Allemand, F. (2004). Spatial working memory deficits in children at ages 3-4 who were low birth weight, preterm infants. *Neuropsychology* 18, 673–678. doi: 10.1037/0894-4105.18.4.673

Wanless, S. B., McClelland, M. M., Lan, X., Son, S. H., Cameron, C. E., Morrison, F. J., et al. (2013). Gender differences in behavioral regulation in four societies: the United States, Taiwan, South, and China. *Early Child. Res. Q.* 28, 621–633. doi: 10.1016/j.ecresq.2013.04.002

Wass, S. V., Scerif, G., and Johnson, M. H. (2012). Training attentional control and working memory – Is younger, better? *Dev. Rev.* 32, 360–387. doi: 10.1016/j.dr.2012.07.001

Welsh, M., and Peterson, E. (2014). Issues in the conceptualization and assessment of hot executive functions in childhood. *J. Int. Neuropsychol. Soc.* 20, 152–156. doi: 10.1017/S1355617713001379

Wiebe, S. A., Espy, K. A., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., and Espy, K. A. (2011). The structure of executive function in 3-year-olds. *J. Exp. Child Psychol.* 108, 436–452. doi: 10.1016/j.jecp.2010.08.008

Willoughby, M., Kupersmidt, J., Voegler-Lee, and Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Dev. Neuropsychol.* 36, 162–180. doi: 10.1080/87565641.2010.549980

Willoughby, M. T., and Blair, C. B. (2011). Test-retest reliability of a new executive function test battery for use in early childhood. *Child Neuropsychol.* 17, 564–579. doi: 10.1080/09297049.2011.554390

Willoughby, M. T., Blair, C. B., Wirth, R. J., and Greenberg, M. (2010). The measurement of executive function at age 3 years: psychometric propertiesand criterion validity of a new battery of tasks. *Psychol. Assess.* 22, 306–317. doi: 10.1037/a0018708

Zelazo, P. D., and Reznick, J. S. (1991). Age-related asynchrony of knowledge and action. *Child Dev.* 62, 719–735. doi: 10.2307/1131173

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 January 2014; accepted: 24 June 2014; published online: 22 July 2014.
*Citation: Mulder H, Hoofs H, Verhagen J, van der Veen I and Leseman PPM (2014) Psychometric properties and convergent and predictive validity of an executive function test battery for two-year-olds. Front. Psychol. **5**:733. doi: 10.3389/fpsyg.2014.00733*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Mulder, Hoofs, Verhagen, van der Veen and Leseman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 16 April 2014
doi: 10.3389/fpsyg.2014.00326

Development of neural mechanisms of conflict and error processing during childhood: implications for self-regulation

Purificación Checa[1,2], M. C. Castellanos[1,2], Alicia Abundis-Gutiérrez[1,2] and M. Rosario Rueda[1,2]*

[1] *Department of Experimental Psychology, Faculty of Psychology, University of Granada, Granada, Spain*
[2] *Developmental Cognitive Neuroscience Lab, Center for Research on Mind, Brain and Behavior, University of Granada, Granada, Spain*

Edited by:
Philip D. Zelazo, University of Minnesota, USA

Reviewed by:
Sebastian J. Lipina, Unidad de Neurobiología Aplicada, Centro de Educación Médica e Investigaciones Clínicas "Norberto Quirno"–Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina
Nicolas Chevalier, University of Edinburgh, UK

***Correspondence:**
M. Rosario Rueda, Department of Experimental Psychology and Center for Research on Mind, Brain and Behavior, University of Granada, Campus de Cartuja s/n, 18071 Granada, Spain
e-mail: rorueda@ugr.es

Regulation of thoughts and behavior requires attention, particularly when there is conflict between alternative responses or when errors are to be prevented or corrected. Conflict monitoring and error processing are functions of the executive attention network, a neurocognitive system that greatly matures during childhood. In this study, we examined the development of brain mechanisms underlying conflict and error processing with event-related potentials (ERPs), and explored the relationship between brain function and individual differences in the ability to self-regulate behavior. Three groups of children aged 4–6, 7–9, and 10–13 years, and a group of adults performed a child-friendly version of the flanker task while ERPs were registered. Marked developmental changes were observed in both conflict processing and brain reactions to errors. After controlling by age, higher self-regulation skills are associated with smaller amplitude of the conflict effect but greater amplitude of the error-related negativity. Additionally, we found that electrophysiological measures of conflict and error monitoring predict individual differences in impulsivity and the capacity to delay gratification. These findings inform of brain mechanisms underlying the development of cognitive control and self-regulation.

Keywords: executive attention, error processing, conflict resolution, self-regulation, development

INTRODUCTION

Regulating behavior is effortful and requires attention particularly when relying on automatic well-learned actions is insufficient or impossible. Automatic behavior is not appropriate when alternative responses are available and the dominant more automatic response is not the desired one. In such situations errors are likely and detecting them also requires attention (Posner and DiGirolamo, 1998). Error-detection and conflict monitoring are mechanisms related to executive control and have been associated with activation of the executive attention network (EAN), a neural network involving the anterior cingulate cortex (ACC), the anterior insula, and other regions of the prefrontal cortex that are well connected with the basal ganglia and the autonomic nervous system (Posner et al., 2007). Thus, the EAN plays an important role in the regulation of thoughts and emotions (Rueda et al., 2011).

In the laboratory, conflict tasks such as the Flanker or Stroop-like tasks are used to measure executive control processes involving the EAN. Participants are slower and less accurate to respond to trials entailing conflict, as when distracting stimulation surrounds the target (flanker interference effect; Eriksen and Eriksen, 1974). Using this type of task with electrophysiological recordings allows studying the brain mechanisms related to executive control. Many studies have examined modulation of brain's event-related potentials (ERPs) by conflict and have consistently shown that conflict modulates the amplitude of a negative deflection that appears around 200 to 450 ms after presentation of the target (N200, often also called N450; Liotti et al., 2000; Hanslmayr et al., 2008; Szucs and Soltész, 2012). This effect is distributed over mid frontal channels and has been related to activation originated in the ACC (van Veen and Carter, 2002).

Another ERP index associated with action regulation is the error-related negativity (ERN; Luu et al., 2003). The ERN is a negative deflection that appears around 100 ms after the commission of an error (Gehring et al., 1993). A widely accepted account of the ERN suggests that it reflects conflict at the response selection level, signaling a mismatch between the representation of the correct response and the one finally produced (Carter et al., 1999). The conflict monitoring account of ERN predicts activation of the EAN when detecting errors, and in fact both conflict monitoring and the ERN appear to have common cognitive mechanisms and shared neural basis (Yeung et al., 2004).

A second potential also modulated by the commission of an error is a positivity (P_e) that arises around 200–300 ms after the response. This component is considered a later error-related signal, which reflects accumulated evidence that an error has been committed (Steinhauser and Yeung, 2010). The P_e has been associated with awareness of the commission of an error (Kaiser et al., 1997; O'Connell et al., 2007; Shalgi et al., 2009) and with the emotional significance of the error (Leuthold and Sommer, 1999; Ridderinkhof et al., 2009). The rostral part of ACC, a structure associated with self-referential thinking, is involved in the generation of P_e (Herrmann et al., 2004).

Over the course of development children become increasingly able to deal with conflict, showing a major development of this ability during the preschool years (Rueda et al., 2004a; Huizinga et al., 2006; Garon et al., 2008; Lee et al., 2013). Using conflict tasks adapted to children, it has been shown that young children (under age 7 years) show larger conflict effects compared to older children and adults (Rueda et al., 2005a). However, additional data with other tasks involving executive control indicate that this function shows a protracted development during childhood and depending on the demands of executive processes may extend to adolescence and early adulthood (Davidson et al., 2006; Waszak et al., 2010).

Electrophysiological studies have reported changes in brain activity during performance of conflict tasks with age. As adults, children show larger amplitude in trials involving conflict in ERP components around the expected latencies. However, compared to adults, conflict effects in children are larger in amplitude and duration, and have a more anterior distribution (Rueda et al., 2004b; Abundis-Guitiérrez et al., 2014). Moreover, the N450 effect decreases in amplitude with age, which some authors have interpreted as an index of improvement in efficiency of the EAN (Jonkman, 2006; Lamm et al., 2006; Espinet et al., 2012). Evidence from fMRI studies indicates that poorer performance on conflict task in children, compared to adults, relates to their ability to effectively recruit areas involved in cognitive control, such as the ventro-lateral prefrontal cortex (Bunge et al., 2002; Durston et al., 2002; Konrad et al., 2005).

Other studies have investigated the development of error processing during childhood. Errors can be caused by a premature execution of the response, and are often regarded as an instance of impulsive action (Botvinick et al., 2001; Pailing et al., 2002). This idea is supported by the fact that the reaction time (RT) in erroneous responses is usually faster than the RT in correct responses. Compared to adults, children show larger RT difference between correct and error responses (Davies et al., 2004a; Wiersema et al., 2007), indicating that children are more impulsive than adults, likely related to their greater difficulty in inhibiting inappropriate responses.

There is evidence that the ERN is present in children as young as 5 years of age when simple tasks are employed (Torpey et al., 2009). However, studies using more complex tasks have demonstrated that ERN is not clearly shown by children until late childhood (Davies et al., 2004a; Wiersema et al., 2007) or even until early adulthood (Ladouceur et al., 2007). Moreover, whereas the amplitude of the ERN has been positively correlated with age, the P_e appears to be more invariant across development than the ERN (Hajcak and Foti, 2008). Some studies have reported P_e effects of similar amplitude for children and adults (Davies et al., 2004a; Wiersema et al., 2007).

Over and above the existence of an ontogenetic developmental trajectory for the ability to regulate behavior, individuals show large differences in their self-regulatory capacities. Individual differences in regulation have been broadly studied in temperament research. Three broad dimensions characterize temperament during childhood and adolescence (Rothbart and Bates, 2006; Rothbart, 2007), namely: extraversion/surgency (E/S), negative affectivity (NA), and effortful control (EC). The first two dimensions describe individual differences in approaching and avoiding reactivity, respectively, whereas the third dimension describes individual differences in the ability to regulate emotions and actions in an internally guided or voluntary mode. EC is thus the temperament dimension most closely linked to the concept of self-regulation. Also, executive control mechanisms (i.e., conflict processing and error detection) have been conceptually and empirically linked to EC (Rueda, 2012). Many studies have shown an association between performance of conflict tasks and parent- or self- reported measures of EC (Gerardi-Caulton, 2000; Rothbart et al., 2003; Checa et al., 2008). Likewise, individual differences in conflict processing have been related to emotional regulation. It has been reported that children who obtain lower conflict scores show reduced tendency to frustration (Gerardi, 1997), less negative emotional reactions (Gonzalez et al., 2001), and better emotional regulation when facing challenging social situations such as receiving an undesired gift (Simonds et al., 2007). Moreover, low EC has been associated with disruptive behavior and poor sociability in school (Checa et al., 2008), presence of externalizing (Valiente et al., 2003; Olson et al., 2005; Eisenberg et al., 2009) and internalizing (Oldehinkel et al., 2004) behavior problems, and symptoms of depression (Verstraeten et al., 2009).

The current study had two main goals. First, we aimed at exploring the development of neural mechanisms related to conflict and error processing from early to late childhood. The second goal was to further examine the relationship between individual differences in functional efficiency of the EAN and behavioral and temperamental measures of self-regulation. For that purpose, children between 4 and 13 years of age and adults were asked to perform a child friendly flanker task while electrophysiological activity was recorded. The task was designed as to allow studying separately brain activation related to target and response processing. By using this procedure we intended to measure both the ERP related to conflict and error processing. Additionally, children's self-regulatory skills were measured using a delay of gratification task and parent-reported temperament questionnaires. We expected a decrease in the size of the conflict-related potential as a function of age, primarily between preschoolers and older children. In addition, if larger amplitude on the conflict-related potential indexes poorer efficiency of the EAN, this effect should also be negatively related to behavioral self-regulation abilities. Finally, we expected to observe developmental changes in error processing from the preschool period to late childhood, and anticipated a positive relationship between efficiency of neural mechanisms related to error detection and children's self-regulatory skills.

MATERIALS AND METHODS

PARTICIPANTS

A total of 20 adults (14 women; mean age = 23.6 years; SD = 2.6 years) and 47 children participated in the study. Children were divided into three groups: 4–6 year olds ($n = 17$, 10 girls; mean age = 5 years, SD = 1.04 years), 7–9 year olds ($n = 15$, 6 girls; mean age = 8.25 years; SD = 1 year), and 10–13 year olds ($n = 15$; 7 girls; mean age = 10.8 years, SD = 1.44 years). All participants were from an urban area of southern Spain, and had a similar social background. Information on mother's educational

level was collected for the sample of children according to a scale ranging from primary studies (1) to university degree (5). The average scores for children of the different age groups were 4.69 (SD = 0.11), 4.85 (SD = 0.10) and 4.87 (SD = 0.11), respectively, for 4–6, 7–9, and 10–13 year olds, which did not differ significantly from each other ($F < 1$). The parents of the children were contacted by phone and invited to participate in the study. They were part of a database of families who participated in prior studies and expressed their wiliness to participate in future studies. The adults were students of the University of Granada who signed up to participate in the study through the website of the department. The study protocol and recruitment procedures were approved by the Ethics Board of the University of Granada in accord with the Spanish Ministry of Science and Innovation norms for research involving humans. Participation was voluntary, and both the children's caregivers and the adults gave written consent.

PROCEDURE

Participants first completed the Flanker task while their brain activation was registered using a high-density (128-channels) electroencephalography (EEG) system. Fitting the sensor-net on, checking impedances, and completing the computer task took about 35 min, including brief breaks between blocks of trials. Once this task was completed, the sensor net was taken-off, and participants completed the self-report (adults) or parent-report (children's caregivers) version of the temperament questionnaire, which took about 15 min. Finally, children completed a delay of gratification task. All participants performed the different tasks in the same order. At the end of the experimental session, a T-shirt of the lab and other small presents were offered to the children in appreciation for their collaboration. Adults received course credits in accordance to the norms of the Department of Experimental Psychology of the University of Granada.

EXPERIMENTAL TASK

We designed a child-friendly flanker task using pictures of round and square robots as stimuli (see **Figure 1**). Each trial started with a fixation cross displayed at the center of the screen for a variable duration, randomly selected between 600 and 1200 ms. Subsequently, a cartoon picture of a row of five robots was presented either above or below the fixation cross. Participants were asked to focus on the robot in the middle and indicate whether it was round or square by pressing the corresponding button. The robot shape-to-response button mapping was counterbalanced across participants. Robots on the sides could be of the same (congruent) or different (incongruent) shape as that of the middle robot. Flanking robots were congruent in half of the trials, and the congruency condition was randomly selected for each trial. The response could be made during presentation of the target or up to 800 ms after it disappeared. The duration of the target was adjusted in each trial according to the participant's performance in the previous trial. When an error was made, the response was omitted or given off time, the target duration was increased by 50 ms in the following trial. Alternatively, the target duration in trial n + 1 was decreased by 50 ms when the response in trial n was correct. Using this procedure, we intended to adjust the difficulty of the task across participants of different ages, as well as obtaining a significant number of errors in order to examine error potentials. Following the response, a 600 ms-lasting feedback was provided. The feedback consisted of a visual animation of the central figure plus an auditory word ("yes" for correct response, "no" for incorrect response, and "late" for omission or off-time responses). Participants completed 192 trials divided in eight blocks with small breaks between blocks.

TEMPERAMENT QUESTIONNAIRES

The short form of the parent-report version of the Children's Behavioral Questionnaire (CBQ; Putnam and Rothbart, 2006)

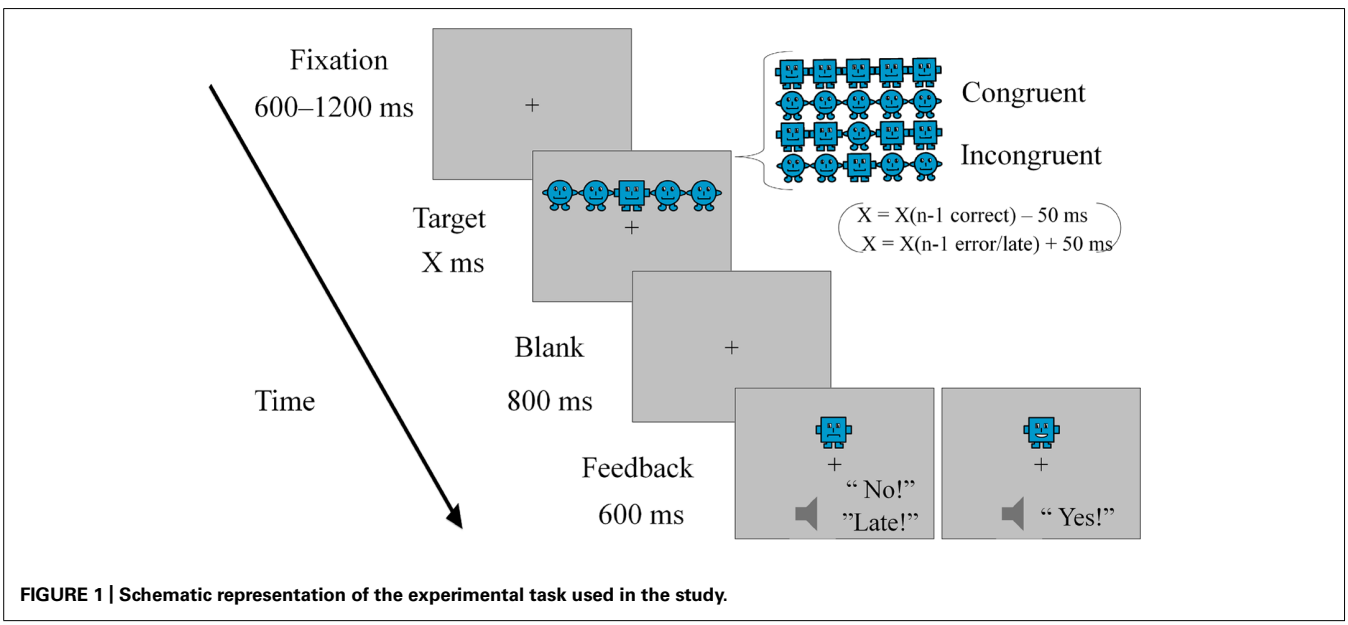


FIGURE 1 | Schematic representation of the experimental task used in the study.

was used for children between 4 and 8 years of age, whereas the Early Adolescence Temperament Questionnaire – Revised (EATQ-R; Ellis and Rothbart, 2001) was used for 9–13 years old children, and the Adult Temperament Questionnaire (ATQ; Rothbart et al., 2000) was used for adults. These questionnaires consist of a number of questions about people's reactions in daily life situations that can be grouped into three main factors: EC, SU, and NA. The ATQ also includes a factor of orienting sensitivity (OS). The internal reliability for each factor in our sample was: Cronbach's $\alpha = 0.60$ for EC, $\alpha = 0.77$ for SU, and $\alpha = 0.86$ for NA in the CBQ; $\alpha = 0.85$ for EC, $\alpha = 0.90$ for AF, $\alpha = 0.34$ for SU, and $\alpha = 0.63$ for NA in the EATQ-R; and $\alpha = 0.53$ for EC; $\alpha = 0.50$ for SU; $\alpha = 0.65$ for NA, and $\alpha = 0.82$ for OS in the ATQ. Only the factors with $\alpha > 0.50$ were included in subsequent analyses.

DELAY OF GRATIFICATION TASK

We used a modified version of Thompson et al. (1997) Delay of Gratification task. We included six types of trial, which were created by crossing three types of reward (stickers, 5 cents of euro coins, and candies) and two types of choice: delay for oneself (DS) or delay for another person (DO). In the first condition (DS), children chose between obtaining: (a) a present for themselves immediately or (b) two presents for themselves at the end of the task. In the DO condition, children chose between obtaining: (a) a present for themselves immediately or (b) a present for themselves and a present for the experimenter at the end of the task. Each participant made 12 choices, 6 of each type. The dependent variable for this task was the percentage of delay choices.

EEG RECORDING AND DATA PROCESSING

Electroencephalography was recorded using the 128-channel Geodesic Sensor Net (EGI Software: www.egi.com). Impedances for each channel were measured prior to recording and monitored during the EEG session. Channels with impedances exceeding 50 kΩ at recording were noted and discarded for further processing. The EEG signal was digitized at 250 Hz and 0.1–100 Hz band pass-filtered during the recording (time constant of 9 s). Recording in every channel was vertex referenced. After recording, data were filtered using a 0.3–12 Hz band pass filter. Continuous data were segmented in various ways in order to examine brain activation locked to different events: target and response. The epochs were 900 ms long (−200 to 700 ms) for target-locked ERPs and 1000 ms long (−600 to 400 ms) for response-locked ERPs. In both cases, we used the 200 ms prior to the event as baseline.

Segmented files were scanned for artifacts with the artifact detection tool provided by the EGI software Net Station. We used a threshold of 100 μV for eye blink or eye movements. Segments containing eye blinks or movements as well as segments with more than 25 bad channels were rejected. Data for each trial were also visually inspected to make sure the parameters of the artifact detection tool were appropriate for each participant. Individual ERPs data were included in the analyses as long as they had a minimum of 12 clean segments per experimental condition. The selection criterion was reached by 50 participants: 12 children in the 4–6 year group (7 girls, mean age = 5.1 years, SD = 0.9); 14 children in the 7–9 year group (6 girls, mean age = 8.1 years, SD = 0.93); 10 children in the 10–13 year group (3 girls, mean age = 11 years, SD = 1.1), and 14 adults (9 women, mean age = 26.5 years, SD = 5.3).

RESULTS

BEHAVIORAL RESULTS

Reaction time and accuracy data per age group in various conditions of the experimental task are presented in **Table 1**. Median RTs per experimental condition was used to measure speed of responses and percentage of errors (both errors of commission and omission) to measure accuracy. As shown in **Figure 2C**, the percentage of errors committed in the task was about 20% for all age groups, which provided sufficient error responses as to examine the brain reaction to errors and different types of feedback.

Separate 4 (Age Group) × 2 (Flanker Type) ANOVAs with median RTs and percentage of errors as dependent measures were conducted. For RT, results revealed a significant main effect of Age Group, $F(3,46) = 42.32$, $p < 0.001$. Planned contrasts revealed that adults were faster than all children groups; [$F(1,46) = 116$, $p < 0.001$; $F(1,46) = 30$, $p < 0.001$ and $F(1,46) = 4.5$, $p < 0.05$ comparisons with 4–6, 7–9, and 10–13 years old, respectively]. Also, the 10–13 years olds were faster than the 7–9 year group, $F(1,46) = 8.3$, $p < 0.01$; and the 10–13 and 7–9 years old groups were faster than the 4–6 year group, $F(1,46) = 62$, $p < 0.001$ and $F(1,46) = 31$, $p < 0.001$, respectively (see **Figure 2A**). The main effect of Flanker Type was also significant, $F(1,46) = 41$, $p < 0.001$, indicating faster responses in congruent compared to incongruent trials. The Age Group × Flanker Type interaction

Table 1 | Performance of the experimental task.

	% Errors				RT (ms)			
	OV-Com.	OV-Om.	Cong.	Incon.	Cong.	Incon.	Correct resp.	Errors resp.
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Adults	18 (5.5)	4.8 (1.1)	12.9 (6.0)	23.1 (7.2)	384 (31)	408 (31)	393 (29)	356 (29)
10–13 year	20.8 (2.9)	1.5 (1.2)	17.7 (4.3)	24 (4.8)	500 (59)	527 (57)	513 (58)	433 (70)
7–9 year	17.7 (5.3)	2.4 (2.5)	14.9 (4.7)	20.4 (6.9)	656 (131)	694 (134)	676 (137)	574 (123)
4–6 year	20.6 (5.4)	4.1 (3.3)	18.5 (6.0)	22.7 (7.0)	940 (205)	1000 (252)	965 (200)	794 (184)

OV-Com., Overall commission errors; OV-Om., Overall Omission errors; M, Mean; SD, Standard Deviation; Cong., Congruent trials; Incon., Incongruent trials.

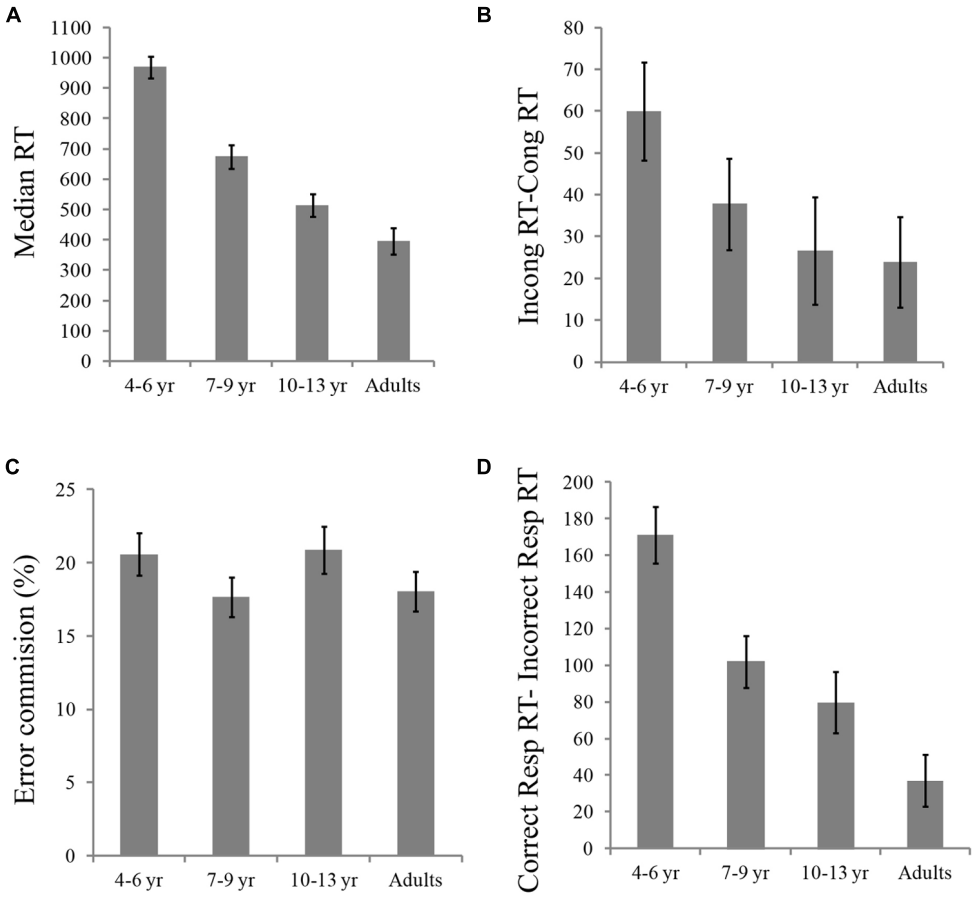


FIGURE 2 | (A) Median RTs (ms) in each Age Group; **(B)** Flanker effect (RTs in incongruent – RTs in congruent trials) by age group in ms; **(C)** Percentage of commission of errors by age group; and **(D)** Impulsivity index (RTs in correct responses – RTs in incorrect responses) by age group, in ms.

was not significant, however, planned comparisons showed that the flanker interference effect (incongruent vs. congruent RT) was significantly larger for the 4–6 year group than for adults, $F(1,46) = 5.12$, $p < 0.05$, and marginally larger for the 4–6 year group compared to the 10–13 years old, $F(1,46) = 3.7$, $p = 0.06$ (see **Figure 2B**).

Using the percentage of errors as dependent variable, we found a significant main effect of flanker type, $F(1,46) = 47.2$; $p < 0.001$, indicating smaller percentage of errors in congruent compared to incongruent trials. Neither the main effect of Age Group, $p > 0.5$, nor the Age Group × Flanker Type interaction, $p > 0.05$, were significant on the accuracy analysis (see **Figure 2C**).

We also examined differences in RT for correct compared to incorrect responses across age groups. Overall, participants were faster when their responses were incorrect compared to correct, $F(1,46) = 165$, $p < 0.001$. The effect of Response Type interacted with age, $F(3,46) = 14.16$, $p < 0.001$. Planned contrasts indicated that the incorrect vs. correct difference in RT was smaller for adults compared to the 4–6 year, $F(1,46) = 41.26$, $p < 0.001$, 7–9 year, $F(1,46) = 10.5$, $p < 0.01$, and 10–13 year, $F(1,46) = 3.7$, $p < 0.05$, groups. Also, this difference was larger for the 4–6 year group than for the 7–9 and 10–13 year groups, $F(1,46) = 10.9$, $p < 0.01$ and $F(1,46) = 16.1$, $p < 0.001$, respectively, whereas there were no differences ($F > 1$) between children in the 7–9 and 10–13 year groups (see **Figure 2D**).

DELAY OF GRATIFICATION

Percentages of delay choices obtained in the DoG task were entered in a 3 (Age Group) × 2 (Delay Type: self vs. other) ANOVA. Results revealed a significant main effect of Age Group, $F(2,43) = 20.2$, $p < 0.001$. Planned comparisons indicated that 7–9 years olds (65%) and 10–13 year olds (79.4%) did not differ on the percentage of delay choices. However, the percentage of delay choices was smaller for the 4–6 year group (27%) compared to the 7–9 year group, $F(2,43) = 19.6$, $p < 0.001$, and the 10–13 year group, $F(1,43) = 37$, $p < 0.001$. The main effect of Delay Type was also significant, $F(1,43) = 5.9$, $p < 0.05$, with larger percentage of delay choices for oneself (62.9%) than for someone else (51.8%). The Age Group × Delay Type interaction was not significant, $p > 0.05$.

ERPs RESULTS

Target-locked ERPs

Averaged ERPs per Flanker Type condition and Age Group are presented in **Figure 3A**. **Figure 3B** illustrates the topographic

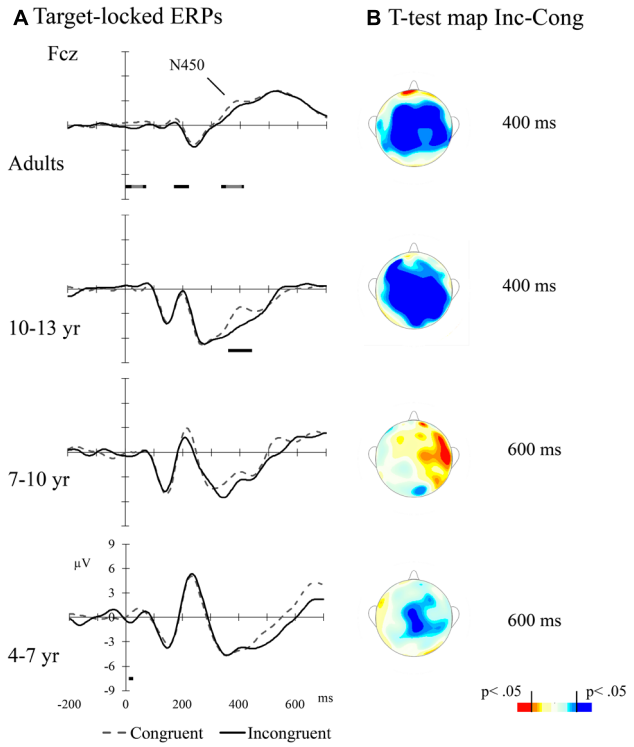


FIGURE 3 | (A) Target-locked ERPs for adults and children at mid-frontal leads. The bars above the temporal scale show when the Error-Correct response *t*-test is significant (light gray: $p < 0.01$, black: $p < 0.05$, dark gray: $p < 0.1$); **(B)** Scalp distributions of incongruent – congruent contrasts at particular times after target presentation (*t*-test; 400 ms for adults and for 10–13 years old groups, and 600 ms for 7–10 year. old and 4–7 year old groups)

distribution of incongruent minus congruent difference (congruency effect) at times of interest. The amplitude difference between congruent and incongruent trials appears to be largest between 350 and 450 ms for adults and older children, and some delayed for younger groups. In order to analyze the congruency effect in the different Age Groups, the mean amplitude per condition was calculated at different time windows: 350–450 ms post-target for adults and 10–13 year group, and 550–650 ms post-target for 7–10 and 4–7 year groups of children. Data from two lead positions over the midline, Cz, and Fcz, were included in this analysis. Thus, a 4 (Age Group) × 2 (Flanker Type) × 2 (electrode position: anterior-Fcz and posterior-Cz) ANOVA was run using the mean amplitude for the time windows specified above as dependent variable. The main effects of Age Group, $F(3,46) = 4.04$, $p < 0.05$, and Flanker Type, $F(1,46) = 12.12$, $p < 0.01$, were significant. The second indicating that the amplitude was more negative for incongruent compared to congruent trials. The Age Group × Flanker Type interaction was not significant ($p > 0.1$). The main effect of Electrode Position was also significant, $F(1,46) = 126.95$, $p < 0.001$, with larger amplitude at Fcz than at Cz. This effect was qualified by a significant Age Group × Electrode Position interaction, $F(3,46) = 4.43$, $p < 0.001$, showing that the Fcz vs. Cz amplitude difference was smaller in adults than in the 10–13 year, $F(1,46) = 7.78$, $p < 0.001$, and the 7–9 year, $F(1,46) = 6.12$, $p < 0.05$, groups. Also, it was smaller for the 4–6 year group compared to 10–13 year, $F(1,46) = 7.12$, $p < 0.05$, and 7–9 year, $F(1,46) = 5.50$, $p < 0.05$, groups. There was no difference between adults and 4–6 year group ($p > 0.1$), and between 10–13 year and 7–10 year groups ($p > 0.1$).

Response-locked ERPs

Table 2 shows mean amplitudes per condition and Age Group in the various ERP components of interest (i.e., N450, ERN, and Pe). Also, averaged ERPs for correct vs. error responses and the different age groups are presented in **Figure 4A**. **Figure 4B** illustrates the topographic distribution of the error minus correct responses difference at time points corresponding to the ERN and P_e peaks.

A 4 (Age Group) × 2 (Response Type: correct vs. error) × 2 (Electrode Position: Fcz and Cz) ANOVA was run using a residualized ERN as dependent variable (see **Table 2**). This measure was calculated using linear regression to partial out the variability from the ERN amplitude due to the preceding positivity (see Santesso et al., 2005; Santesso and Segalowitz, 2008). The VD of the linear regression was the peak amplitude of the ERN at the time window from 0 to 100 ms post-response, and the VI was the peak amplitude of the preceding positivity

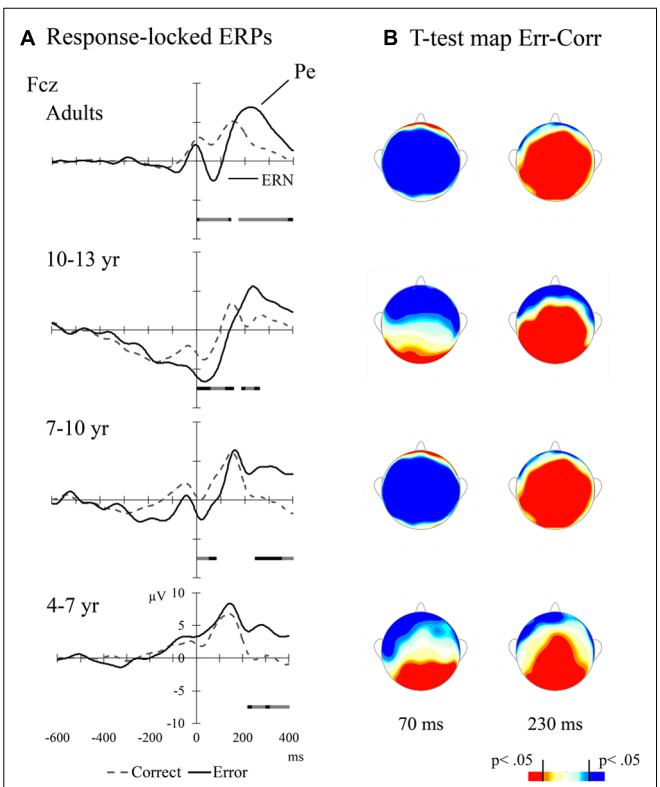


FIGURE 4 | (A) Response-locked ERPs for adults and children at mid-frontal leads. The bars above the temporal scale show when the Error-Correct response *t*-test is significant (light gray: $p < 0.01$, black: $p < 0.05$, dark gray: $p < 0.1$); **(B)** Scalp distributions of the error vs. correct responses *t*-test values at particular times after the response (70 ms for ERN and 230 ms for P_e in all age groups)

Table 2 | Amplitude of ERPs components by channel, group, and conditions.

		N450			ERN		P_e	
		Fcz	Cz		Fcz	Cz	Fcz	Cz
Group	Cond.	M (SD)	M (SD)	Cond.	M (SD)	M (SD)	M (SD)	M (SD)
Adults	Cong.	2.9 (2.4)	5.7 (3.1)	Co	1.6 (2.8)	4.7 (3.2)	5.4 (2.7)	7.5 (3.3)
	Incon.	2.3 (2.1)	2.8 (0.7)	Err	−2.8 (3.1)	0.4 (2.5)	7.7 (3.5)	8.8 (3.6)
10–13 year	Cong.	−2.8 (4.3)	2.9 (2.9)	Co	−4.1 (4.7)	3.5 (3.4)	4.4 (2.6)	8.7 (3.3)
	Incon.	−4.5 (4.2)	3.6 (1.1)	Err	−7.8 (6.0)	1.4 (5.1)	6.7 (5.2)	14.2 (3.8)
7–9 year	Cong.	1.4 (4.8)	6.2 (4.4)	Co	−0.2 (5.2)	4.3 (3.1)	6.8 (4.6)	7.5 (4.5)
	Incon.	0.4 (5.9)	5.1 (1.4)	Err	−3.7 (6.6)	2.4 (5.9)	7.6 (8.1)	11.9 (9.7)
4–6 year	Cong.	2.0 (4.3)	5.0 (3.1)	Co	1.3 (6.8)	2.2 (5.7)	7.2 (5.6)	5.2 (5.7)
	Incon.	0.1 (4.5)	3.8 (1.1)	Err	1.2 (5.4)	2.6 (4.9)	9.4 (6.7)	9.0 (6.4)

The amplitude values are expressed in μVolts. M, Mean; SD, Standard Deviation; Cong., Congruent trials; Incon., Incongruent trials; Co., Correct responses; Err., Erroneous responses.

at the −100 to 0 ms pre-response time window, and a residual score was saved. We found significant main effects of Response Type, $F(1,46) = 29.61$, $p < 0.001$, with larger negative amplitude for errors compared to correct responses; and Electrode Position, $F(1,46) = 171.39$, $p < 0.001$, with larger amplitude at Fcz than Cz. Both Response Type and Electrode Position interacted with Age Group, $F(3,46) = 21.98$, $p < 0.001$ and $F(3,46) = 21.98$, $p < 0.001$, respectively. The difference in amplitude between error and correct responses was significant in adults, $F(1,46) = 27.04$, $p < 0.001$, in 10–13 year group, $F(1,46) = 8.41$, $p < 0.01$, and 7–9 year children, $F(1,46) = 10.31$, $p < 0.01$, but not in 4–6 year children, $p > 0.1$. The Fcz amplitude was larger than Cz amplitude in adults, $F(1,46) = 24.52$, $p < 0.001$, 10–13 year, $F(1,46) = 121.58$, $p < 0.001$, and 7–9 year children, $F(1,46) = 66.98$, $p < 0.001$, but not in 4–6 year children, $p > 0.1$. The interaction Response Type × Electrode Position was significant, $F(1,46) = 7.98$, $p < 0.01$, because the error-correct response difference in amplitude was larger at Fcz than at Cz, $F(1,46) = 7.21$, $p < 0.01$.

Additionally, all age groups showed later larger positive amplitudes for error compared to correct responses (P_e effect; see **Table 2**). In order to analyze this effect, peak amplitudes per response type were calculated in a time window ranging from 130 to 270 ms post-response for each participant, and included in a 4 (Age Group) × 2 (Response Type) × 2 (Electrode Position: Fcz and Cz) ANOVA. The Response Type main effect was significant, $F(1,46) = 24.77$, $p < 0.001$, with larger positive amplitude for error than for correct responses. The Electrode Position main effect was significant, $F(1,46) = 30.54$, $p < 0.001$, with larger amplitude at Cz than Fcz. This effect was mediated by Age Group, $F(1,46) = 12.21$, $p < 0.001$, showing that the amplitude was larger at Cz than Fcz for adults, $F(1,46) = 4.64$, $p = 0.036$, 10–13 year $F(1,46) = 45.12$, $p < 0.001$, 7–9 year, $F(1,46) = 11.30$, $p = 0.002$, but not for 4–6 year children, $p > 0.1$. The Response Type × Electrode Position interaction was also significant, $F(1,46) = 19.12$, $p < 0.001$. This interaction was mediated by Age Group, $F(1,46) = 6.56$, $p < 0.001$. Planned comparisons indicated that the difference in P_e amplitude between errors and correct responses at Fcz was marginally significant in adults, $F(1,46) = 3.94$, $p = 0.053$, older children, $F(1,46) = 3.06$, $p = 0.087$, and younger children, $F(1,46) = 3.25$, $p = 0.078$ groups, but not in the medium children group, $p > 0.1$. At Cz, the P_e amplitude was larger for errors than correct responses in children groups [for older, $F(1,46) = 16.85$, $p < 0.001$; for medium, $F(1,46) = 15.33$, $p < 0.001$, for younger, $F(1,46) = 9.53$, $p < 0.001$] but not in adults, $p > 0.1$.

Correlations

Two scores of flanker interference (i.e., incongruent vs. congruent flankers) were obtained for each participant in both RT (FI_{RT}) and percentage of errors (FI_{ERR}). We also obtained an index of impulsivity (IM) by subtracting the median RTs for error responses from the median RTs for correct responses. Correlation between these scores and data on DoG and temperament showed that FI_{ERR} was positively correlated with impulsivity, $r = 0.52$, $p < 0.05$; $r = 0.47$, $p < 0.05$ after controlling by age. The percentage of delayed choices at the DoG task did not correlate with neither of flanker interference and impulsivity scores. The correlation between the FI_{ERR} and EC was significant in adults, $r = -0.58$, $p < 0.05$; and children younger than 9 years EC, $r = -0.72$, $p < 0.001$. FI_{RT} was also correlated with NA, $r = 0.51$, $p < 0.05$, and SU, $r = 0.46$, $p < 0.05$, for children younger than 9 years of age. The percentage of delayed choices at the DoG task did not correlate with any of the temperamental factors.

Additionally, we calculated ERP indexes of conflict (N450) and error processing (ERN and P_e). The N450 index was obtained by subtracting the mean amplitude for congruent trials from the mean amplitude for incongruent trials at Fcz at the following time windows: 350–450 ms post-target for adults and 10–13 year children and 550–650 ms post-target for 7–9 year and 4–6 year children groups. The ERN index was calculated subtracting the residualized ERN amplitude for correct responses from the residualized ERN amplitude for incorrect responses at Fcz. The P_e index was calculated by subtracting the peak amplitude for correct responses

from the peak amplitude of the error responses at time window at of 130–270 ms post-response at Cz. Pearson correlation between those ERP scores and indexes of task performance are presented in **Table 3**. We found a positive correlation between the N450 index and IM score, $r = 0.47$, $p < 0.01$; $r = 0.49$, $p < 0.01$ after controlling by age. Also, the N450 index was negatively related to the percentage of delayed choices in the DS condition of the DoG task, $r = -0.36$, $p < 0.05$ after controlling by age. These correlations are plotted in **Figure 5**. Additionally, the P_e index was correlated with IM, $r = -0.27$, $p = 0.05$, and with temperamental factor of SU, $r = -0.32$, $p < 0.05$; $r = -0.27$, $p < 0.05$, after controlling by age. Finally, significant correlation was also found between the ERN index and FI_{RT}, $r = -0.26$, $p < 0.05$, as well as with the total of percentage of delay choices in the DoG task after controlling by age, $r = 0.31$, $p = 0.05$.

DISCUSSION

The aim of the current study was to investigate the neural mechanisms of executive attention and examine their relation to the development of self-regulation from early to late childhood. To asses executive attention we used a child-friendly flanker task designed to measure conflict resolution, as well as error and feedback processing. In this task, the duration of the target was adjusted in a trial-by-trial basis for each participant in order to ensure an equivalent level of task difficulty for participants of different ages.

DEVELOPMENT OF CONFLICT AND ERROR PROCESSING

Behavioral results of our study showed poorer executive control skills in children of the youngest group (4–6 year olds) compared to older children and adults. Despite performing the experimental task at equivalent accuracy levels, the youngest group showed larger flanker interference score and larger impulsivity index than adults and older children (see **Figure 2**). Moreover, young children showed a significant smaller capacity to delay gratification compared to 7–10 and 10–13 year olds. All three measures suggest the existence of a major developmental change between preschool ages and middle-to-late childhood. This result is generally consistent with data from other developmental studies using a variety of tasks targeting executive functions, which also indicate that early to middle childhood constitutes an important developmental period of this function. This is for instance the case in studies using the dimensional card sorting task (Zelazo et al., 1996) inhibitory control (Bedard et al., 2002), and flanker tasks (Rueda et al., 2004a). Likewise, Wiersema et al. (2007) and Davies et al. (2004b) also found a decrease in the error vs. correct response time differences with age.

In addition to the behavioral level of analysis, we were able to study the neural basis of executive attention by registering electrophysiological patterns of activations during task performance. In our study, manipulation of the congruency of flankers modulated the amplitude of the target-locked N450 potential. This modulation was clearly observed in adults and 10–13 year olds in a group of frontally distributed channels (see **Figure 3**). In younger children, this modulation appeared to emerge later and to be sustained longer, although, as revealed by *t*-tests analyses, did not reach significance. Data in the literature about developmental changes in conflict-related modulations of target-evoked mid-frontal potentials greatly depend on the task being used.

Table 3 | Pearson correlations between electrophysiological indexes and indexes of performance of the experimental task.

		FI_{RT}	FI_{ERR}	IM	DOG
N450	Fcz	0.14 (0.11)	0.05 (−0.06)	**0.47** (0.49**)**	−0.34 **(−0.36*)**
ERN	Fcz	**−0.26*** (−0.21)	0.08 (−0.06)	−0.05 (0.02)	0.20 **(0.31*)**
P_e	Cz	0.14 (0.07)	−0.17 (−0.10)	**0.27*** (0.15)	0.15 (0.08)

*Data between parentheses are correlations controlled by age. Significance values: **p < 0.01; *p < 0.05.*

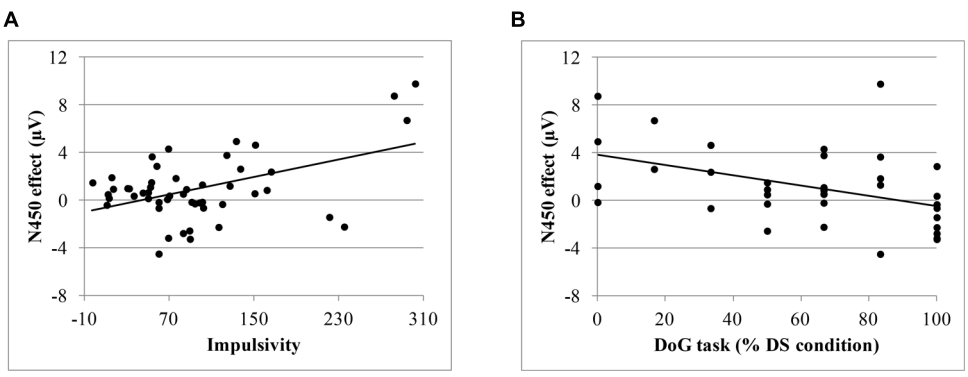


FIGURE 5 | (A) Correlation between the N450 effect and the impulsivity score; **(B)** Correlation between the N450 effect and the percentage of delayed choices in the delay for oneself (DS) condition of the DoG task.

Several studies using Go–NoGo tasks have reported larger conflict effects in the N200/N450 amplitude by young children compared to older children and adults (Lamm et al., 2006; Hämmerer et al., 2010). This result suggests that the larger the effect on the amplitude of the N200/N450 the poorer the executive control efficiency. As a matter of fact, Lamm et al. (2006) reported an age-related decrease in N200/N450 amplitude between 7 and 16 years of age. However, using a flanker task with arrows, Ladouceur et al. (2007) found that only late adolescents (i.e., older than 14 years) and adults showed larger N200 amplitude in trials with incongruent flankers, while an early adolescents group also included in the study did not show the effect. Our results are consistent with data from this study as well as with those reported by Rueda et al. (2004b) where young children did not show negative amplitude modulations by flanker congruency but a sustained frontal effect after 500 ms post-target. Generally, the longer delay and duration of the effect in younger ages may, at least partially, explain young children's poorer functional efficiency of the EAN.

Regarding neural processes of error monitoring, we found clear differences in the developmental trajectories of the ERN and P_e components. All age groups showed a clear P_e component. However, the ERN was not observed in 4–6 year old children. This result is consistent with prior data on the development of error processing during childhood (Davies et al., 2004b; Wiersema et al., 2007). There is evidence suggesting that the ERN consist of an early, and probably subconscious, signal of mismatch between the represented goal and the response being produced (Yeung et al., 2004). On the other hand, the P_e component appears to reflect accumulated evidence that an error was committed and the negative evaluation associated with it (Ridderinkhof et al., 2009; Steinhauser and Yeung, 2010). One possible interpretation is that the detection of errors in young children might depend to a greater extent on affective processes (an evaluation of the response and the negative outcome of this evaluation). Such processes would be slower than the subconscious mismatch thought to give rise to the ERN, and might involve the ventral (more affective) division of the ACC. In support of differential underlying mechanisms, some studies using dipole modeling have shown that the ERN and the P_e are generated in different brain regions (van Veen and Carter, 2002; Herrmann et al., 2004). Generally, both the dorsal and ventral vision of the ACC have been involved in executive control, the ventral division being particularly important in situations that are emotionally relevant (Bush et al., 2000). The ventral ACC facilitates executive control in situations signaled by emotion (Kanske and Kotz, 2011). Since each ACC division is associated with different cognitive mechanisms different developmental trajectories might be expected. Children in our study showed adult-like brain responses in the latency of the P_e component, a result that suggests that the ventral executive control system shows an earlier maturational trajectory than the cognitive dorsal system.

CONFLICT AND ERROR PROCESSING AND SELF-REGULATION

The second goal of this study was to investigate the relation between the efficiency of EAN and the development of self-regulation. It has been suggested that mechanisms of executive attention are key to the development of self-regulatory skills (Rueda et al., 2011). Executive attention and the temperamental factor of EC are closely related concepts that depict different levels of analysis (i.e., cognitive and behavioral, respectively) of the ability to regulate behavior (Gerardi, 1997; Gonzalez et al., 2001; Simonds et al., 2007; Checa et al., 2008). Results of our study support the connection between cognitive measures of executive attention and the temperament factor of EC. Moreover, behavioral self-regulation measures and efficiency of EAN were related in our data. We found a correlation between higher impulsivity and poorer capacity to delay gratification and amplitude of the N450. As discussed above, larger N450 conflict effect is associated poorer executive attention efficiency. Thus, children showing poorer efficiency of the system at the neural level also show poorer regulatory skills at the behavioral level. Importantly, this result is obtained after age differences in the different measures are controlled for. These findings complement prior work supporting the existence of a link between efficiency of the EAN and individual differences in the ability to regulate actions (Posner and Rothbart, 1998; Rueda et al., 2011).

Previous research had linked impulsivity to difficulties in inhibitory control (Patterson and Newman, 1993; Barkley, 1997; Gonzalez et al., 2001; Enticott et al., 2006; Spinrad et al., 2012). Our data also reveal a positive correlation between the ability to inhibit inappropriate responses and impulsivity as well as a positive correlation between amplitude of the ERN and the ability to delay gratificaction. Previous studies have also shown a link between amplitude of the ERN and self- as well as social-regulation capacities (Santesso and Segalowitz, 2009). Moreover, individual differences in impulsivity were also associated with amplitude of the P_e component. These data are in line with previous studies showing that individuals who exhibited more impulsive behaviors displayed poor EAN efficiency, using the same or similar indexes of impulsivity as the one used in the present research (Pailing et al., 2002; Ruchsow et al., 2005), as well as using self-reported measures of impulsivity (Gonzalez et al., 2001; Heritage and Benning, 2013). In support of this relationship, there is evidence that children with ADHD, a disorder associated with impulsive behavior, show less efficiency in neural mechanism and structures within the EAN (Liotti et al., 2005; Wiersema et al., 2005; van Meel et al., 2007). All this evidence indicates that weaker and slower reactions related with conflict and error processing in frontal brain regions underlay behavioral patterns characterized by poor self-regulatory capacity. This conclusion is consistent with the role of the EAN in the Posner's model (Posner et al., 2007).

According to Posner and Rothbart (2007), the EAN is involved in the regulation of emotional reactivity, both negative and positive. Our results are aligned with this idea. We found that higher flanker interference scores, indicative of poorer attentional control, were positively related to both negative affectivity as well as surgency. The use of attentional control to regulate emotions is thought to be supported by the system related to attentional selectivity (i.e., orienting network) in the early years, and relying on the developing EAN later on (Rothbart et al., 2011). The EAN is involved in controlling affect-related information through its connections with subcortical limbic structures such as the amygdala (Ochsner and Gross, 2004). Previous studies have associated dysfunctions of EAN with the inability to regulate emotions (Gehring

et al., 2000; Ruchsow et al., 2005; Hajcak and Foti, 2008). Moreover, recent evidence shows that reduced activity in areas within EAN is associated with negative affect (Crocker et al., 2012).

Our data also revealed a negative correlation between surgency and the amplitude of the P_e potential. This suggests that excessive positive affect can impair some aspects of error processing. Several studies have reported that positive affect is associated with decreased planning abilities, task switching and worse inhibition abilities (Phillips et al., 2002; Mitchell and Phillips, 2007). We suggest that the development of the EAN, and subsequent enhanced attentional control, provides the attentional flexibility required to regulate approaching tendencies and resists temptations. This is particularly important when current conditions call for actions that conflict with future goals and those action are to be inhibited. The efficiency of EAN to control both positive/approaching as well as negative/avoiding tendencies is important for a broad range of aspects of children's life such as morality and social adjustment (Kochanska et al., 2009), school readiness and academic performance (Checa et al., 2008; Checa and Rueda, 2011; Kim et al., 2013), and behavioral problems (Oldehinkel et al., 2004; Verstraeten et al., 2009).

The relation between efficiency of the EAN and regulation of approaching tendencies is not only restricted to reactive systems of temperament in our data. The ability to resist temptation in favor of long-term goals shows a positive relationship with amplitude of the ERN over and above age (see **Table 3**). There is evidence that success in the DoG task depends on the ability to regulate the attention during the waiting period (Mischel, 1974; Mischel et al., 1989). Additionally, imaging studies have shown that top-down control regions of the prefrontal cortex are activated during the delay period in DoG tasks (Casey et al., 2011; Heatherton and Wagner, 2011). Our data also show that children who were more able to delay gratification were the ones who better recruited the EAN during conflict resolution and error processing. Prior research has shown that performance of the DoG task in childhood predicts the efficiency with which the same individuals perform a Go/No-go task as adolescents and young adults (Eigsti et al., 2006).

CONCLUSION

Data from this study inform about the development of diverse aspects of executive attention and self-regulation. Data from different domains (i.e., cognitive, temperament, and brain function) were taken into account. Results added to the evidence indicating that executive attention shows a period of major development during preschool years (Rueda et al., 2005b). By registering ERPs during performance of a flanker task, we were able to examine neural mechanisms related to conflict and error processing, and found that individual differences in efficiency of those mechanisms predict children's ability to delay gratification and individual differences in impulsivity. Concretely, better error detection predicts larger percentage of delay choices and less impulsive behavior, whereas greater brain commitment (measured with amplitude of the N450 effect) in resolving conflict from incongruent flankers predicts smaller percentages of delay choices and more impulsive responses.

The scope of the current study was limited to the use of one particular experimental paradigm to explore brain mechanisms related to conflict processing and error detection. Flanker tasks are widely used in the literature to examine executive control, and the task utilized in our study had the advantage of adjusting the difficulty to the performance level of each participant; however, replicating the results of the study with other tasks (e.g., Stroop, Go–NoGo) would be desirable. Future studies might also benefit from using longitudinal designs in order to examine individual differences in the developmental trajectory of executive attention.

In sum, data from our study provide evidence that children showing a more efficient engagement of the EAN during development also show better self-regulation skills. In the recent past, mounting evidence is showing that neural mechanisms of executive attention can be enhanced by means of cognitive training (Rueda et al., 2005b). Interventions of this sort have the potential to also enhance children's regulatory skills. For instance, we recently found that children trained in executive attention show better performance in a delay of gratification task compared to untrained peers (Rueda et al., 2012). Self-regulation is key to socialization and academic success (Rueda et al., 2010), therefore understanding the mechanisms underlying the development of this system as well as finding the best ways to boost its efficiency will be matters of great interest in future research.

ACKNOWLEDGMENTS

Research presented in this article was supported by a grant from the Spanish Ministry of Science and Innovation (ref. PSI2011.27746) to M. Rosario Rueda and a pre-doctoral FPU fellowship from the Spanish Ministry of Science and Innovation awarded to the fist author. The research presented in this paper was part of the doctoral dissertation of the first author.

REFERENCES

Abundis-Guitiérrez, A., Checa, P., Castellanos, M. C., and Rueda, M. R. (2014). Electrophysiological correlates of attention networks in childhood and early adulthood. *Neuropsychologia* doi: 10.1016/j.neuropsychologia.2014.02.013 [Epub ahead of print].

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* 121, 65–94. doi: 10.1037/0033-2909.121.1.65

Bedard, A.-C., Nichols, S., Barbosa, J. A., Schachar, R., Logan, G. D., and Tannock, R. (2002). The development of selective inhibitory control across the life span. *Dev. Neuropsychol.* 21, 93–111. doi: 10.1207/S15326942DN2101_5

Botvinick, M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi: 10.1037/0033-295X.108.3.624

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., and Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron* 33, 1–20. doi: 10.1016/S0896-6273(01)00583-9

Bush, G., Luu, P., and Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* 4, 215–222. doi: 10.1016/S1364-6613(00)01483-2

Carter, C. S., Botvinick, M. M., and Cohen, J. D. (1999). The contribution of the anterior cingulate cortex to executive processes in cognition. *Rev. Neurosci.* 10, 49–57. doi: 10.1515/REVNEURO.1999.10.1.49

Casey, B. J., Somerville, L., Gotlib, I., Ayduk, O., Franklin, N., Askren, M., et al. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proc. Natl. Acad. Sci. U.S.A.* 36, 14998–15003. doi: 10.1073/pnas.1108561108

Checa, P., Rodriguez-Bailon, R., and Rueda, M. R. (2008). Neurocognitive and temperamental systems of self-regulation and early adolescents' school competence. *Mind Brain Educ.* 2, 177–187. doi: 10.1111/j.1751-228X.2008.00052.x

Checa, P., and Rueda, M. R. (2011). Behavioral and brain measures of executive attention and school competence in late childhood. *Dev. Neuropsychol.* 36, 1018–1032. doi: 10.1080/87565641.2011.591857

Crocker, L., Heller, W., Spielberg, J., Warren, S., Bredemeier, K., Sutton, B., et al. (2012). Neural mechanisms of attentional control differentiate trait and state negative affect. *Front. Psychol.* 3:298. doi: 10.3389/fpsyg.2012.00298

Davidson, M. C., Amso, D., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006

Davies, P. L., Segalowitz, S. J., and Gavin, W. J. (2004a). Development of error-monitoring event-related potentials in adolescents. *Ann. N. Y. Acad. Sci.* 1021, 324–328. doi: 10.1196/annals.1308.039

Davies, P. L., Segalowitz, S. J., and Gavin, W. J. (2004b). Development of response monitoring ERPs in 7- to 25- year-old. *Dev. Neuropsychol.* 25, 355–376. doi: 10.1207/s15326942dn2503_6

Durston, S., Thomas, K. M., Yang, Y., Ulug, A. M., Zimmerman, R. D., and Casey, B. (2002). A neural basis for the development of inhibitory control. *Dev. Sci.* 5, F9–F16. doi: 10.1111/1467-7687.00235

Eigsti, I. M., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M., et al. (2006). Predicting cognitive control from preschool to late adolescence and young adulthood. *Psychol. Sci.* 17, 478–484. doi: 10.1111/j.1467-9280.2006.01732.x

Eisenberg, N., Valiente, C., Spinrad, T. L., Liew, J., Zhou, Q., Losoya, S. H., et al. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to their externalizing, internalizing, and co-occurring behavior problems. *Dev. Psychol.* 45, 988–1008. doi: 10.1037/a0016213

Ellis, L. K., and Rothbart, M. K. (2001). "Revision of the early adolescent temperament questionnaire," in *Poster presented at the 2001 Biennal Meeting of the Society for Research in Child Development*, Minneapolis, MN.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Espinet, S., Anderson, J., and Zelazo, P. (2012). N2 amplitude as a neural marker of executive funtion in young children: an ERP study of children who switch versus perseverate on the Dimensional Change Card Sort. *Dev. Cogn. Neurosci.* 2(Suppl.1), S49–S58. doi: 10.1016/j.dcn.2011.12.002

Enticott, P., Ogloff, J., and Bradshaw, J. (2006). Association between laboratory measures of executive inhibitory control and self-reported impulsivity. *Pers. Individ. Diff.* 41, 285–296. doi: 10.1016/j.paid.2006.01.011

Garon, N., Bryson, S., and Smith, I. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi: 10.1037/0033-2909.134.1.31

Gehring, W. J., Gross, B., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390. doi: 10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., Himle, J., and Nisenson, L. G. (2000). Action-monitoring dysfunction in obssessive-compulsive disorder. *Psychol. Sci.* 11, 1–5. doi: 10.1111/1467-9280.00206

Gerardi, G. (1997). *Development of Executive Attention and Self-regulation in the third Year of Life.* Dissertation presented to the Department of Psychology, University of Oregon, Eugene, OR.

Gerardi-Caulton, G. (2000). Sensitivity to spatial conflict and the development of self-regulation in children 24-36 months of age. *Dev. Sci.* 3, 397–404. doi: 10.1111/1467-7687.00134

Gonzalez, C., Fuentes, L. J., Carranza, J. A., and Estevez, A. F. (2001). Temperament and attention in the self-regulation of 7-year-old children. *Pers. Individ. Diff.* 30, 931–946. doi: 10.1016/S0191-8869(00)00084-2

Hajcak, G., and Foti, D. (2008). Errors are aversive: defensive motivation and the error-related negativity. *Psychol. Sci.* 19, 103–108. doi: 10.1111/j.1467-9280.2008.02053.x

Hämmerer, D., Li, S.-C., Müller, V., and Lindenberger, U. (2010). An electrophysiological study of response conflict processing across the lifespan: assessing the roles of conflict monitoring, cue utilization, response anticipation, and response suppression. *Neuropsychologia* 48, 3305–3316. doi: 10.1016/j.neuropsychologia.2010.07.014

Hanslmayr, S., Pastotter, B., Bauml, K. H., Gruber, S., Wimber, M., and Klimesch, W. (2008). The electrophysiological dynamics of interference during the stroop task. *J. Cogn. Neurosci.* 20, 215–225. doi: 10.1162/jocn.2008.20.2.215

Heatherton, T. F., and Wagner, D. D. (2011). Cognitive neuroscience of self-regulation failure. *Trends Cogn. Sci.* 15, 132–139. doi: 10.1016/j.tics.2010.12.005

Heritage, A., and Benning, S. (2013). Impulsivity and response modulation deficit in psychopathy: evidence from ERN and N1. *J. Abnorm. Psychol.* 122, 215–222. doi: 10.1037/a0030039

Herrmann, M. J., Rommler, J., Ehlis, A. C., Heidrich, A., and Fallgatter, A. J. (2004). Source localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe). *Cogn. Brain Res.* 20, 294–299. doi: 10.1016/j.cogbrainres.2004.02.013

Huizinga, M., Dolan, C., and van der Molen, M. (2006). Age-related change in executive funtion: development trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Jonkman, L. M. (2006). The development of preparation, conflict monitoring and inhibition from early childhood to young adulthood: a go / nogo ERP study. *Brain Res.* 1097, 181–193. doi: 10.1016/j.brainres.2006.04.064

Kaiser, J., Barker, R., Haenschel, C., Baldeweg, T., and Gruzelier, J. H. (1997). Hypnosis and event-related potential correlates of error processing in a Stroop-like paradigm: a test of the frontal hypothesis. *Int. J. Psychophysiol.* 27, 215–222. doi: 10.1016/S0167-8760(97)00055-x

Kanske, P., and Kotz, S. A. (2011). Emotion speeds up conflict resolution: a new role for the ventral anterior cingulate cortex? *Cereb. Cortex* 21, 911–919. doi: 10.1093/cercor/bhq157

Kim, S., Nordling, J. K., Yoon, J. E., Boldt, L. J., and Kochanska, G. (2013). Effortful control in "hot" and "cold" tasks differentially predicts children's behavioral problems and academic outcomes. *J. Abnorm. Child Psychol.* 41, 43–56. doi: 10.1007/s10802-012-9661-4

Kochanska, G., Barry, R. A., Jimenez, N. B., Hollatz, A. L., and Woodard, J. (2009). Guilt and effortful control: two mechanisms that prevent disruptive developmental trajectories. *J. Pers. Soc. Psychol.* 97, 322–333. doi: 10.1037/a0015471

Konrad, K., Neufang, S., Thiel, C. M., Specht, K., Hanisch, C., Fan, P., et al. (2005). Development of attentional networks: an fMRI study with children and adults. *Neuroimage* 28, 429–439. doi: 10.1016/j.neuroimage.2005.06.065

Ladouceur, C. D., Dahl, R. E., and Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: ERP and source localization. *Dev. Sci.* 10, 874–891. doi: 10.1111/j.1467-7687.2007.00639.x

Lamm, C., Zelazo, P. D., and Lewis, M. D. (2006). Neural correlates of cognitive control in childhood and adolescence: disentangling the contributions of age and executive function. *Neuropsychologia* 44, 2139–2148. doi: 10.1016/j.neuropsychologia.2005.10.013

Lee, K., Bull, R., and Ho, R. (2013). Developmental changes in executive funtioning. *Child Dev.* 84, 1933–1953. doi: 10.1111/cdev.12096

Leuthold, H., and Sommer, W. (1999). ERP correlates of error processing in spatial S-R compatibility tasks. *Clin. Neurophysiol.* 110, 342–357. doi: 10.1016/S1388-2457(98)00058-3

Liotti, M., Pliszka, S. R., Perez, R., Kothmann, D., and Woldorff, M. G. (2005). Abnormal brain activity related to performance monitoring and error detection in children with ADHD. *Cortex* 41, 377–388. doi: 10.1016/S0010-9452(08)70274-0

Liotti, M., Woldorff, M. G., Perez, R., and Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia* 38, 701–711. doi: 10.1016/S0028-3932(99)00106-2

Luu, P., Tucker, D. M., Derryberry, D., Reed, M., and Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychol. Sci.* 14, 47–53. doi: 10.1111/1467-9280.01417

Mischel, W. (1974). "Processes in delay of gratification," in *Advances in Experimental Social Psychology*, Vol. 7, ed. L. Berkowitz (San Diego, CA: Academic Press), 249–292.

Mischel, W., Shoda, Y., and Rodriguez, M. L. (1989). Delay of gratification in children, *Science* 244, 933–938. doi: 10.1126/science.2658056

Mitchell, R. L. C., and Phillips, L. H., (2007). The psychological, neurochemical and functional neuroanatomical mediators of the effects of positive and negative mood on executive functions. *Neuropsychologia* 45, 617–629. doi: 10.1016/j.neuropsychologia.2006.06.030

Ochsner, K. N., and Gross, J. J. (2004). *Thinking Makes It So: A Social Cognitive Neuroscience Approach to Emotion Regulation*. New York, NY: Guilford Press.

O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H., et al. (2007). The role of cingulate cortex in the detection of errors with and without awareness: a high-density electrical mapping study. *Eur. J. Neurosci.* 25, 2571–2579. doi: 10.1111/j.1460-9568.2007.05477.x

Oldehinkel, A. J., Hartman, C. A., De Winter, A. F., Veenstra, R., and Ormel, J. (2004). Temperament profiles associated with internalizing and externalizing problems in preadolescence. *Dev. Psychopathol.* 16, 421–440. doi: 10.1017/S0954579404044591

Olson, S. L., Sameroff, A. J., Kerr, D. C., Lopez, N. L., and Wellman, H. M. (2005). Developmental foundations of externalizing problems in young children: the role of effortful control. *Dev. Psychopathol.* 17, 25–45. doi: 10.1017/S0954579405050029

Pailing, P. E., Segalowitz, S. J., Dywan, J., and Davies, P. L. (2002). Error negativity and response control. *Psychophysiology* 39, 198–206. doi: 10.1111/1469-8986.3920198

Patterson, C. M., and Newman, J. P. (1993). Reflectivity and learning from aversive events: toward a psychological mechanism for the syndromes of disinhibition. *Psychol. Rev.* 100, 716–736. doi: 10.1037/ 0033-295X.100.4.716

Phillips, L. H., Bull, R., Adams, E., and Fraser, L. (2002). Positive mood and executive function: evidence from Stroop, and fluency tasks. *Emotion* 2, 21–22. doi: 10.1037/1528-3542.2.1.12

Posner, M. I., and DiGirolamo, G. J. (1998). *Executive Attention: Conflict, Target Detection, and Cognitive Control*. Cambridge, MA: The MIT Press.

Posner, M. I., and Rothbart, M. K. (1998). Attention, self-regulation and consciousness. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 353, 1915–1927. doi: 10.1098/rstb.1998.0344

Posner, M. I., and Rothbart, M. K. (2007). Research on attention networks as a model for the integration on psychological science. *Annu. Rev. Psychol.* 58, 1–23. doi: 10.1146/annurev.psych.58.110405.085516

Posner, M. I., Rothbart, M. K., Sheese, B., and Tang, Y. (2007). The anterior cingulate gyrus and the mechanism of self-regulation. *Cogn. Affect. Behav. Neurosci.* 7, 391–395. doi: 10.3758/CABN.7.4.391

Putnam, S. P., and Rothbart, M. K. (2006). Development of short and very short forms of the children's behavior questionnaire. *J. Pers. Assess.* 87, 102–112. doi: 10.1207/s15327752jpa8701_09

Ridderinkhof, K. R., Ramautar, J., and Wijnen, J. G. (2009). To P_E or not to P_E: A P3-like ERP component reflecting the processing of response errors. *Psychophysiology* 46, 531–538. doi: 10.1111/j.1469-8986.2009.00790.x

Rothbart, M. K. (2007). Temperament, development, and personality. *Curr. Dir. Psychol. Sci.* 16, 207–212. doi: 10.1111/j.1467-8721.2007.00505.x

Rothbart, M. K., Ahadi, S., and Evans, D. (2000). Temperament and personality: origins and outcomes. *J. Pers. Soc. Psychol.* 78, 122–135. doi: 10.1037//0022-3514.78.1.122

Rothbart, M. K., and Bates, J. E. (2006). "Temperament in children' development," in *Handbook of Child Psychology,* Vol. 3, *Social, emotional, and personality development,* eds R. Lerner, N. Eisenberg, and W. Damon (New York: Wiley), 99–166.

Rothbart, M. K., Ellis, L. K., Rueda, M. R., and Posner, M. I. (2003). Developing mechanisms of conflict resolution. *J. Pers.* 71, 1113–1143. doi: 10.1111/1467-6494.7106009

Rothbart, M. K., Sheese, B. E., Rueda, M. R., and Posner, M. I. (2011). Developing mechanisms of self-regulation in early life. *Emot. Rev.* 3, 207–213. doi: 10.1177/1754073910387943

Ruchsow, M., Spitzer. M., Gron, G., Grothe, J., and Kiefer, M. (2005). Error processing and impulsiveness in normals: evidence from event related potentials. *Cogn. Brain Res.* 24, 317–325. doi: 10.1016/j.cogbrainres.2005.02.003

Rueda, M. R. (2012). "Effortful control," in *Handbook of Temperament*, Chap. 8, eds M. Zentner and R. Shiner (New York: Guilford Press), 145–167.

Rueda, M. R., Checa, P., and Combita, L. (2012). Enhanced efficiency of the executive attention network after training in preschool children: immediate changes and effects after two months. *Dev. Cogn. Neurosci.* 2(Suppl.1), S192–S204. doi: 10.1016/j.dcn.2011.09.004

Rueda, M. R., Checa, P., and Rothbart, M. K. (2010). Contributions of attentional control to socioemotional and academic development. *Early Educ. Dev.* 21, 744–764. doi: 10.1080/10409289.2010.510055

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004a). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Rueda, M. R., Posner, M. I., Rothbart, M. K., and Davis-Stober, C. P. (2004b). Development of the time course for processing conflict: an event-related potentials study with 4 year olds and adults. *BMC Neurosci.* 5:39. doi: 10.1186/1471-2202-5-39

Rueda, M., Posner, M. I., and Rothbart, M. K. (2005a). The development of executive attention: contributions to the emergence of self-regulation. *Dev. Neuropsychol.* 28, 573–594. doi: 10.1207/s15326942dn2802_2

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., and Posner, M. I. (2005b). Training, maturation and genetic influences on the development of executive attention. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14931–14936. doi: 10.1073/pnas.0506897102

Rueda, M. R., Posner, M. I., and Rothbart, M. K. (2011). "Attentional control and self-regulation," in *Handbook of Self-Regulation: Research, Theory and Applications*, 2nd Edn, eds K. D. Vohs and R. F. Baumeister (New York: The Guilford Press), 284–299.

Santesso, D. L., and Segalowitz, S. J. (2008). Developmental differences in error-related ERPs in middle- to late-adolescent males. *Dev. Psychol.* 1, 205–217. doi: 10.1037/0012-1649.44.1.205

Santesso, D. L., Segalowitz, S. J., and Schmidt, L. A. (2005). ERP correlates of error monitoring in 10-year olds are related to socialization. *Biol. Psychol.* 70, 79–87. doi: 10.1016/j.biopsycho.2004.12.004

Santesso, D., and Segalowitz, S. J. (2009). The error-related negativity is related to risk taking and empathy in young men. *Psychophysiology* 46, 143–152. doi: 10.1111/j.1469-8986.2008.00714.x

Shalgi, S., Barkan, I., and Deouell, L. Y. (2009). On the positive side of error processing: error-awareness positivity revisited. *Eur. J. Neurosci.* 29, 1522–1532. doi: 10.1111/j.1460-9568.2009.06690.x

Simonds, J., Kieras, J. E., Rueda, M., and Rothbart, M. K. (2007). Effortful control, executive attention, and emotional regulation in 7-10-year-old children. *Cogn. Dev.* 22, 474–488. doi: 10.1016/j.cogdev.2007.08.009

Spinrad, T. L., Eisenberg, N., Kassondra, M. S., Eggum, N. D., Reiser, M., Edwards, A., et al. (2012). Longitudinal relations among maternal behaviors, effortful control and young children's committed compliance. *Dev. Psychol.* 48, 552–556. doi: 10.1037/a0025898

Steinhauser, M., and Yeung, N. (2010). Decision processes in human performance monitoring. *J. Neurosci.* 30, 15643–15653. doi: 10.1523/JNEUROSCI.1899-10.2010

Szucs, D., and Soltész, F. (2012). Functional definition of the N450 event-related brain potential marker of conflict processing: a numerical Stroop study. *BMC Neurosci.* 3:35. doi: 10.1186/1471-2202-13-35

Thompson, C., Barresi, J., and Moore, C. (1997). The development of future-oriented prudence and altruism in preschoolers. *Cogn. Dev.* 12, 199–212. doi: 10.1016/S0885-2014(97)90013-7

Torpey, D. C., Hajcak, G., and Klein, D. N. (2009). An examination of error-related brain activity and its modulation by error value in young children. *Dev. Neuropsychol.* 34, 749–761. doi: 10.1080/87565640903265103

van Meel, C. S., Heslenfeld, D. J., and Sergeant, J. A. (2007). Adaptive control deficits in attention-deficit/hyperactivity disorder (ADHD): the role of error processing. *Pychiatry Res.* 151, 211–220. doi: 10.1016/j.psychres.2006.05.011

van Veen, V., and Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Psychol. Behav.* 77, 477–482. doi: 10.1016/S0031-9384(02)00930-7

Valiente, C., Eisenberg, N., Smith, C. L., Reiser, M., Fabes, R. A., Losoya, S., et al. (2003). The relations of effortful control and reactive control to children 's externalizing problems: a longitudinal assessment. *J. Pers.* 71, 1171–1196. doi: 10.1111/1467-6494.7106011

Verstraeten, K., Vasey, M. W., Raes, F., and Bijttebier, P. (2009). Temperament and risk for depressive symptoms in adolescence: mediation by rumination and moderation by effortful control. *J. Abnorm. Child Psychol.* 37, 349–361. doi: 10.1007/s10802-008-9293-x

Waszak, F., Li, S. C., and Hommel B. (2010). The development of attentional networks: cross-sectional findings from a life span sample. *Dev. Psychol.* 46, 337–349. doi: 10.1037/a0018541

Wiersema, J. R., van der Meere, J. J., and Roeyers, H. (2005). ERP correlates of impaired error monitoring in children with ADHD. *J. Neural. Transm.* 112, 1417–1430. doi: 10.1007/s00702-005-0276-6

Wiersema, J. R., van der Meere, J. J., and Roeyers, H. (2007). Developmental changes in error monitoring: an event-related potential study. *Neuropsychologia* 45, 1649–1657. doi: 10.1016/j.neuropsychologia.2007.01.004

Yeung, N., Botvinick, M. M., and Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* 111, 931–959. doi: 10.1037/0033-295X.111

Zelazo, P. D., Burack, J. A., Benedetto, E., and Frye, D. (1996). Theory of mind and rule use in individuals with Down's syndrome: a test of the uniqueness and specificity claims. *J. Child Psychol. Psychiatr.* 37, 479–484. doi: 10.1111/j.1469-7610.1996.tb01429.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 January 2014; accepted: 28 March 2014; published online: 16 April 2014.
Citation: Checa P, Castellanos MC, Abundis-Gutiérrez A and Rueda MR (2014) Development of neural mechanisms of conflict and error processing during childhood: implications for self-regulation. Front. Psychol. ***5****:326. doi: 10.3389/fpsyg.2014.00326*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Checa, Castellanos, Abundis-Gutiérrez and Rueda. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 22 April 2014
doi: 10.3389/fpsyg.2014.00331

Stability of executive function and predictions to adaptive behavior from middle childhood to pre-adolescence

Madeline B. Harms[1]*, Vivian Zayas[2], Andrew N. Meltzoff[3] and Stephanie M. Carlson[1]

[1] *Institute of Child Development, University of Minnesota, Minneapolis, MN, USA*
[2] *Department of Psychology, Cornell University, Ithaca, NY, USA*
[3] *Institute for Learning and Brain Sciences, University of Washington, Seattle, WA, USA*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Birgit Elsner, Universität Postdam, Germany
Georgiana Susa, Babes-Bolyai University, Romania

****Correspondence:***
Madeline B. Harms, Institute of Child Development, University of Minnesota, 51 East River Rd., Minneapolis, MN 55455, USA
e-mail: harms124@umn.edu

The shift from childhood to adolescence is characterized by rapid remodeling of the brain and increased risk-taking behaviors. Current theories hypothesize that developmental enhancements in sensitivity to affective environmental cues in adolescence may undermine executive function (EF) and increase the likelihood of problematic behaviors. In the current study, we examined the extent to which EF in childhood predicts EF in early adolescence. We also tested whether individual differences in neural responses to affective cues (rewards/punishments) in *childhood* serve as a biological marker for EF, sensation-seeking, academic performance, and social skills in early adolescence. At age 8, 84 children completed a gambling task while event-related potentials (ERPs) were recorded. We examined the extent to which selections resulting in rewards or losses in this task elicited (i) the P300, a post-stimulus waveform reflecting the allocation of attentional resources toward a stimulus, and (ii) the SPN, a pre-stimulus anticipatory waveform reflecting a neural representation of a "hunch" about an outcome that originates in insula and ventromedial PFC. Children also completed a Dimensional Change Card-Sort (DCCS) and Flanker task to measure EF. At age 12, 78 children repeated the DCCS and Flanker and completed a battery of questionnaires. Flanker and DCCS accuracy at age 8 predicted Flanker and DCCS performance at age 12, respectively. Individual differences in the magnitude of P300 (to losses vs. rewards) and SPN (preceding outcomes with a high probability of punishment) at age 8 predicted self-reported sensation seeking (lower) and teacher-rated academic performance (higher) at age 12. We suggest there is stability in EF from age 8 to 12, and that childhood neural sensitivity to reward and punishment predicts individual differences in sensation seeking and adaptive behaviors in children entering adolescence.

Keywords: executive function, affective decision-making, event-related potentials, adolescence, reward processing

INTRODUCTION

Executive function (EF) is comprised of a constellation of functions involving the control of thought and action, including the abilities to inhibit pre-potent responses, flexibly shift attention, and update information in working memory (Miyake et al., 2000; Miyake and Friedman, 2012). Recent literature has distinguished between cool EF, which involves the execution of these processes under relatively neutral conditions, and hot EF, which occurs in emotionally salient contexts that may also require risk and reward processing. Cool and hot aspects of EF show protracted maturation across development and may contribute to real-world behavior in different and/or overlapping ways (Zelazo and Carlson, 2012). The goal of the present work was to examine the influences of hot and cool EF and their neural correlates in childhood on adaptive behavior around the transition to adolescence.

EF is readily measurable during the preschool period, especially between ages 3 and 5, but improved performance in EF tasks is seen well into adolescence (for overview see Carlson et al., 2013). The gradual maturation of EF is likely due to the necessity of prefrontal cortex (PFC) engagement, particularly of the dorsolateral region, to perform these high-level cognitive processes (Bunge and Zelazo, 2006). Children as young as 6 years have been shown to activate the PFC when completing EF tasks, but they show a more diffuse network of activation than adults, which suggests that this network gains efficiency with development (Casey et al., 2000). Structurally, the PFC matures slowly across development; indeed, synaptic pruning of this region does not begin in earnest until adolescence (Casey et al., 2000). Behavioral research suggests that EF skills do not reach their full capacity until early adulthood (Steinberg et al., 2008; Zelazo et al., 2013).

As mentioned, two differentiable but related categories of EF—"cool EF" and "hot EF"—have been proposed based on the level of contextual emotional salience. Experimental tasks have been developed to assess both hot and cool EF. Classic cool EF tasks often involve performing mental operations on neutral stimuli. For example, in the flanker task (Eriksen and Eriksen, 1974; Rueda et al., 2004), individuals identify the direction of a target stimulus (an arrow) that is "flanked" by distracters facing the opposite direction. Likewise, in the Dimensional Change Card

Sort task (DCCS; Zelazo, 2006), individuals sort bivalent stimuli on one dimension and then switch to the other (e.g., sort by color then by shape).

In contrast, Hot EF tasks involve performing mental operations in motivationally salient contexts or on motivationally salient stimuli. For example, in Mischel et al.'s (1989) classic delay of gratification task, children must refrain from eating a tempting treat or ringing a bell that would end the delay period in order to receive a larger reward. Likewise, in affective decision-making or gambling tasks, individuals make decisions about risks and potential rewards. In the Iowa Gambling Task, for example, participants choose among four options on each trial, each of which yields either long-term advantages or disadvantages and either short-term rewards or punishments (Bechara, 2004). This task thus involves learning about the most advantageous option in the context of risks and rewards. Although classified in the same "hot EF" category as delay tasks, gambling tasks recruit different cognitive processes and do not always correlate with delay tasks in children (Hongwanishkul et al., 2005). However, because children beyond the preschool period have little difficulty waiting for a reward, gambling tasks are the most paradigmatic method for examining hot EF in older children and adolescents.

There is some controversy as about the degree to which cool and hot EF tasks rely on overlapping vs. dissociated cognitive and neural processes. Using behavioral evidence, some researchers find associations in performance on cool and hot EF tasks (e.g., Carlson and Moses, 2001) whereas others find dissociations (Hongwanishkul et al., 2005; Smith et al., 2012). In a large sample of preschoolers, Carlson et al. (2014) found support for separate but related Conflict (cool) and Delay (hot) factors in a confirmatory factor analysis. At a neural level, by one account, hot and cool EF tasks rely on the same basic circuitry in PFC, but hot EF tasks are more difficult due to the bottom-up affective factors (primarily from reward-sensitive ventral striatum) that must be overcome (Prencipe et al., 2011, see also Reyna and Zayas, 2014). Another account, based on lesion data, suggests more fundamental differences between hot and cool EF, with the former relying primarily on orbitofrontal cortex (Bechara, 2004) and the latter relying on dorsolateral PFC (Casey et al., 2000).

The tasks that are chosen likely play a role in conflicting findings. For example, affective decision-making in a gambling task requires more updating than a delay of gratification task, and thus gambling tasks may relate more strongly to measures of cool EF (Hongwanishkul et al., 2005). In addition, performance on hot gambling tasks tends to lag significantly behind performance on cool EF tasks (Hooper et al., 2004; Prencipe et al., 2011). This may be due to more delayed maturation of the neural circuitry involved in emotion regulation and/or greater sensitivity to affective cues in younger children.

Regardless of the neural mechanisms involved, individual differences in hot and cool EF tend to be persistent over time. This is especially true in preschoolers, who show a high level of stability in their relative performance on both conflict and delay of gratification EF tasks (Carlson et al., 2004; Hughes and Ensor, 2007). We know less about the stability of individual differences in EF beyond the preschool period, but the extant research suggests that individual differences in EF tend to persist over time. For example, two studies (Eigsti et al., 2006; Casey et al., 2011) have found that the proportion of time preschoolers directed their attention away from rewarding stimuli during a delay-of-gratification task predicted their reaction times in a go-no-go task many years later.

NEURAL CORRELATES OF HOT EF DEVELOPMENT

Due to the limitations of scanning young children, the majority of our knowledge about the neural bases for the development of EF comes from studies examining electrical event-related potentials (ERPs) recorded during EF tasks. One component of interest, the P300, is a stimulus-locked component thought to be generated from frontal and temporal-parietal regions, and to be involved in updating working memory and inhibition (Polich, 2007). The P300 is seen approximately 300 ms post-stimulus in adults, but is delayed to 800–1200 ms post-stimulus in children (Tucker, 1993), suggesting increasing efficiency of EF networks with age.

A classic P300 paradigm is an oddball task, in which participants respond to a rare target among many distracters, but it is well established that the P300 is elicited by EF tasks as well. For example, in a flanker task, the P300 has higher amplitude after incongruent vs. congruent trials, suggesting this component might reflect inhibition of extraneous stimulus processing (Tucker, 1993). The P300 has also been found in the context of a hot EF task in 8-year-old children (Carlson et al., 2009). On a child version of the Iowa Gambling Task, the P300 had a higher amplitude after punishment than after reward trials, and the amplitude difference between loss and reward trials predicted children's performance on the task: Those who showed a more pronounced P300 response to losses vs. rewards learned to avoid disadvantageous and high-frequency punishment choices to a greater extent over the course of the task. In this case, the P300 served as a neural signature of focusing attention on a stimulus that provides important information about whether something should be approached or avoided. In the context of this task, greater sensitivity to punishment led to more avoidance of bad plays, and thus, better performance.

Another component of interest is the stimulus-preceding negativity, or SPN. This component occurs after a response has been made and just before feedback occurs. The SPN has recently been measured in children and seems to occur in the context of reward-based tasks. For example, Stavropoulos and Carver (2013) reported an SPN in 6- to 8-year-old children during a reward-based guessing game. They found larger SPN amplitudes for rewards that were accompanied by a smiling face than those accompanied by a scrambled face, suggesting that social stimuli were perceived as more salient. Although the Stavropoulos and Carver (2013) study involved no punishment, when negative feedback does occur, the SPN tends to be larger prior to receiving negative than positive feedback. This pattern has been found in school-age children for both a probabilistic learning task (Groen et al., 2007) and a gambling task (Carlson et al., 2009). Because research indicates that people are generally more sensitive to punishment than reward (Vaish et al., 2008), these results suggest that the SPN may reflect the emotional salience of an anticipated stimulus. Indeed, the SPN appears to be generated by the insular cortex and may reflect dopaminergic activity there (Bocker et al.,

1994). As with the P300, the SPN may also be a neural signature of learning from feedback (Carlson et al., 2009). Given the probable link between SPN and risk and reward-processing, this component could be a particularly informative neural signature to examine prior to adolescence. Children who are more sensitive to anticipated punishment than anticipated reward as reflected by the SPN might show better adaptive outcomes in adolescence.

INDIVIDUAL DIFFERENCES IN EF AND ADAPTIVE BEHAVIOR IN ADOLESCENCE

Adolescence is a time of significant cortical reorganization, potentially even a sensitive period during which current developmental trajectories can be reinforced or re-directed. Many of the neurobehavioral changes that take place during adolescence may be influenced by changes in hormone levels associated with puberty (Steinberg, 2005). At the same time, there is a dramatic change in the context in which teenagers function. Moving from elementary school to middle and high school involves adapting to new peer groups, increased academic expectations, and increased exposure to high-risk activities. Given the high level of flux in the brain, body, and environment, it is unsurprising that behavior problems often emerge for the first time in adolescence.

Adolescents are more likely than older and younger individuals to engage in risky behavior such as the use of illegal drugs (Substance Abuse and Mental Health Services Administration, 2007) and engagement in unsafe sex (Finer and Henshaw, 2006). However, adolescents do not appear to evaluate the risks or consequences of behavior differently than adults in hypothetical situations (e.g., Beyth-Marom et al., 1993), suggesting that the PFC functions adequately in "cool" contexts that are not emotionally salient. Rather, behavioral differences in adolescence are more marked by the "heat of the moment" when a risky decision is made.

Developmental models characterize human adolescence as a period of increased risk-taking due to immature EF/PFC development, which is not yet up to the challenge of coping with more active reward-processing circuitry (i.e., ventral striatum) (Galvan et al., 2006; Ernst et al., 2009; Steinberg, 2010). It is noteworthy that substantial reorganization of the neural systems underlying EF takes place during adolescence. In the PFC, gray matter reaches peak thickness in early adolescence, and is pruned during the next few years (Paus, 2005). In addition, connectivity between limbic and pre-frontal brain regions increases substantially during adolescence (Eluvathingal et al., 2007). The implications of these changes are that by the end of adolescence, the PFC operates more efficiently and there is greater coordination between pre-frontal and limbic systems.

Behavioral and neuroimaging research suggests that increased risk-taking by adolescents may stem from a greater sensitivity to potential rewards than in other age groups. For example, in affective decision-making tasks such as the Iowa Gambling Task, adolescents are more approach-oriented than are pre-adolescents or adults. One study found that although both adults and adolescents played increasingly more from advantageous decks over the course of the task, only adults decreased their plays from disadvantageous decks (Cauffman et al., 2010). Furthermore, in an fMRI study, Galvan et al. (2006) found that adolescents activated the reward-sensitive nucleus accumbens more than children or young adults during a reward-processing task, and activated the OFC, which is thought to play a regulatory role in risk and reward processing, less than adults. Thus, appetitive reward-sensitive systems may mature earlier in adolescence than regulatory systems, possibly contributing to the observed increase in risky behavior (Ernst et al., 2009).

Despite the reported increase in risky behavior among adolescents, substantial individual differences exist. Better EF skills could be a protective factor that reduces risk-taking behavior in adolescents. Research shows that childhood EF, particularly hot EF, predicts a variety of outcomes in adolescence. Seminal work by Mischel et al. (1989; reviewed in Zayas et al., 2014) re-evaluated high-school students who had completed the delay of gratification task during preschool, and found that individuals who refrained from eating a desirable treat and waited 15 min for a larger reward scored significantly higher on their SATs than those who did not wait, independent of IQ assessed at age 4. In addition, parents rated the adolescents who had delayed as preschoolers higher in social cognitive skills and emotional coping. In further follow-up studies, delay of gratification at age 4 predicted more efficient EF (Eigsti et al., 2006) and less interference on a social-reward version of a go-nogo task at the behavior and neural (fMRI) levels (Casey et al., 2011). Other research has found longitudinal relations between preschool delay of gratification performance and physical health. Children who had settled for a lesser reward at age 4 were 30% more likely to be overweight at age 11 (Seeyave et al., 2009). Thus, the ability to delay gratification in childhood appears to reflect individual differences that influence the development of many aspects of adaptive behavior.

In contrast to the delay of gratification task, there is less longitudinal evidence linking performance on hot affective decision making tasks with later EF and life outcomes. Nonetheless, extant work suggests that affective decision-making tasks could prove useful in evaluating propensities toward risk-taking, particularly in adolescence. Adolescents both engage in more sensation seeking behavior than younger children and adults and make more decisions based on reward rather than punishment feedback (Cauffman et al., 2010; Albert and Steinberg, 2011). These patterns might derive from the same mechanisms, such as dopaminergic activity in the brain (e.g., Ernst et al., 2009). Such mechanisms presumably operate earlier in development as well, in which case it may be possible to assess the risk for future sensation seeking behavior by examining behavioral and neural sensitivity to reward and punishment at earlier ages.

Unlike delay tasks, gambling tasks tap into risk as well as reward processing and may involve larger cognitive demands (Hongwanishkul et al., 2005). Research suggests that the ability to optimize one's gambling strategy develops sometime after the emergence of initial cool EF skills (Hooper et al., 2004; Prencipe et al., 2011). Based on simplified versions of the Iowa Gambling Task, young children seem unable to optimize long-term outcomes, responding only to immediate losses or gains. Not until sometime between middle childhood and early adolescence do children begin to integrate the frequency of gains and losses with long-term consequences (Huizenga et al., 2007; Carlson et al.,

2009). Research also suggests that from childhood to adolescence, individuals become more physiologically sensitive to the *anticipation* of gains and losses, showing larger skin conductance responses *before* choosing frequent loss doors at age 16–18 than at age 10–14 (Crone and van der Molen, 2007). Although children may perform as well as adults at these ages, they still show some differential brain activity; for example, 9–12 year-old children have been shown to activate the anterior cingulate cortex (ACC), which is involved in error monitoring, more than adults on high-risk trials. This finding suggests that the task may be more effortful for them (Van Leijenhorst et al., 2006).

In addition to performance indices, individual differences in neural responses during affective decision-making tasks have also been found to predict later behavioral outcomes. In an ERP study of adolescent monozygotic twins, the P300 effect (amplitude of the P300 in loss vs. gain trials) predicted later alcohol abuse: In each twin pair, one individual began to abuse alcohol in adulthood, and these individuals tended to have had lower amplitude P300 responses to loss trials in early adolescence (Carlson et al., 1999). This study suggests that a blunted P300 effect could be an endophenotype for later high-risk behavior. In the current study, one goal was to extend this finding to examine the extent to which neural responses during affective decision making predict less extreme forms of sensation seeking in a low-risk sample. A related goal, given the literature linking preschool EF to life success, was to assess the extent to which both hot and cool EF predict other adaptive outcomes, such as academic performance and social skills.

PRESENT STUDY

The overarching goal of our research was to characterize individual differences in hot and cool EF that might lead individuals to divergent pathways in adolescence. We re-contacted a cohort of typically-developing children who had been assessed at age 8 on both cool EF measures and a relatively hot EF measure (gambling task) when they were 12 years old and entering adolescence. This longitudinal study had two specific aims: (i) to assess the stability of individual differences in cool EF from middle childhood to early adolescence, an age period that has not yet been the focus of longitudinal research on EF, and (ii) to examine the degree to which cool EF, as well as affective decision-making and its neural correlates at age 8 years (middle childhood), predicted adaptive behavior (academic performance, social skills, and sensation seeking) at age 12 years. We chose Flanker and DCCS tasks to examine cool EF, and a child-friendly gambling task to examine hot EF/affective decision-making, as these are the most paradigmatic and well-supported tasks in the literature to measure these constructs. We hypothesized there would be long-term stability of individual differences in EF. With respect to adaptive behavior, we hypothesized that better performance on an affective decision-making task and/or neural correlates of sensitivity to reward and punishment would predict higher academic achievement and social adjustment and lower sensation-seeking in pre-adolescence. This is the first study, to our knowledge, to examine longitudinal correlates of both cool EF and a hot affective decision-making task in this age group.

MATERIALS AND METHODS

TIME 1

Participants

Eighty-four children who were recruited by telephone from the University of Washington (Seattle) participant database when they were 8 years old completed a series of EF tasks. Here, we report data from 78 children (37 males, 41 females) who participated at both age 8 and age 12. This sample had a mean age of 8 years, 4 months ($SD = 8$ months) at Time 1. Participants were primarily white/non-Hispanic. Maternal education (mode) was a 4-year college degree. Written consent from parents and verbal assent from children was obtained.

Procedure

Participants were tested in a laboratory by a female experimenter. All tasks, other than the Peabody Picture Word Vocabulary Test (PPVT-4), were administered on a computer using E-prime software. An electrode sensor cap (Neuroscan 21-channel) was placed on the child's head while they were seated in front of a computer monitor. A chin rest controlled the distance and alignment to the monitor. During the tasks, participants responded by clicking response-specific buttons on a keyboard. Children completed the following four tasks.

Attention network task (Rueda et al., 2004). On this flanker-type task, participants were shown a row of fish and asked to quickly and accurately indicate whether the *central* fish points to the right or left by a key press. The surrounding "flanker" fish pointed in either the congruent or incongruent direction compared to the central fish (50% of trials each). A spatial cue appeared 150 ms before the preceding the target stimulus (central fish) and was presented in the center, top, or bottom of the screen (48 trials each). The target stimulus always appeared in the center of the screen, 450 ms after the offset of this cue. ITIs varied from 400 to 1600 ms. Participants completed 1 practice block of 24 trials and 4 blocks of 48 trials for data collection. Feedback after each trial was given only in the practice condition. Mean accuracy and median reaction times for congruent and incongruent trials were scored.

Dimensional change card sort (adapted from Zelazo, 2006). This task required participants to shift between sorting stimuli by shape or by color. Participants completed one 40-trial block of practice trials in which only the dominant cue (shape) was presented and four 40-trial blocks of test trials which included 75% dominant (shape) trials and 25% non-dominant (color) trials. Each trial consisted of two target stimuli presented in the upper left (red star) and upper right (blue square) corner of the screen. At the start of the trial, a cue "SHAPE" or "COLOR" appeared in the middle of the screen for 1000 ms, along with a test stimulus directly below it. Participants' task was to match the test stimuli (a red square or blue star) to one of the two target stimuli on the dimension (shape or color) indicated by the cue using a key press. The ITI was 1000 ms during which a gray fixation cross appeared in the middle of the screen. No error feedback was given on any of the test trials. Instructions were presented on the computer screen and described to each participant by a female experimenter. Mean accuracy and median reaction times were scored.

Hungry donkey gambling task (HDT) (Crone and van der Molen, 2004). The objective of this task was to win as many apples for the donkey as possible. Before the game, children were shown a prize bin and told that they could select a prize if they gained more apples than they lost (but all were invited to select a prize at the end). In the task, there were four doors participants could choose to open by pressing the corresponding key, among which long-term gains were crossed with frequency of loss. Door A was disadvantageous over time and yielded frequent small losses (8–12 apples lost on 50% of trials), door B was disadvantageous and yielded infrequent large losses (50 apples on 10% of trials), door C was advantageous but yielded frequent miniscule losses (1–3 apples on 50% of trials), and door D was advantageous and yielded infrequent small losses (10 apples on 10% of trials). Doors A and B yielded a net loss of 10 apples and doors C and D yielded a net gain of 10 apples over the course of the task. Gain and loss information was presented on each trial 500 ms after door selection as a column of red apples crossed out (losses) and a column of green apples (gains). This information remained on the screen for 1000 ms. Overall gains and total number of losses were scored across 280 trials, which were split into 4 blocks of 70 trials. The total number of trials in which a net loss (more apples lost than gained) was incurred was used as a performance measure. For more details on this task, see Carlson et al. (2009).

PPVT-4 (Dunn and Dunn, 2007). Children completed this task for an approximation of their verbal IQ. On each trial, the experimenter said a word and children were asked to indicate the corresponding picture from four options. Age-standardized scores were obtained.

EEG recording

Continuous EEG was recorded from 21 channels during the HDT using a Neuroscan net. Electrodes were placed over the left and right prefrontal (Fp1, Fp2), frontal (F3, F4), inferior frontal (F7, F8), temporal (T7, T8), central (C3, C4), parietal (P3, P4), posterior parietal (P7, P8), occipital (O1, O2), and three midline locations (Fz, Cz, Pz). An electrode placed over the left mastoid was used as the online reference for other channels. A NuAmp 40 Channel Neuroscan amplifier was used with a sampling frequency of 1000 Hz and an online band-pass filter of 0.10–200 Hz. EEG activity was filtered offline using a 30 Hz low-pass filter and re-referenced using an average reference of the right and left mastoid electrodes. Trials contaminated by excessive eye movement or muscle artifacts (150 mV from baseline) were excluded. ERP data from 78 children were included in the analysis. For further details, see Carlson et al. (2009).

ERP analysis

We focused on two ERP components from the HDT, the post-outcome P300 and pre-outcome SPN. We calculated the P300 effect for each participant by subtracting the average amplitude (area under the curve) of trials in which a net loss was incurred from the average amplitude of trials in which a net reward was incurred during the period 300–800 ms post-feedback. We calculated the pre-outcome anticipation of loss effect by subtracting mean voltage for the SPN (−150 ms preceding feedback to +50 ms post-feedback) for high-frequency-punishment door selections (doors A and C) from mean voltage for low-frequency-punishment door selections (doors B and D). Positive numbers, therefore, indicate larger (more negative-going) anticipation effects (see Carlson et al., 2009). A minimum of 20 artifact-free trials for each trial type involved in the calculation was used to calculate P300 and SPN effects. Using difference scores for both these components ensures that signal-to-noise ratios are equated across participants despite individual differences in children's distribution of door choices.

TIME 2

Participants

Families who participated at Time 1 (8 years old) were mailed an invitation to participate in a follow-up study 4 years later, along with questionnaire packets and instructions for completing games online. Of these, 78 families (37 males, 41 females) sent back child and parent questionnaires. The mean age of our Time-2 sample was 12 years, 4 months ($SD = 9$ months) and in 6th or 7th grade (both grades are in middle school in the Seattle area). 66 children (31 males, 35 females) sent back teacher-completed questionnaires, and 67 (32 males, 35 females) completed the online games.

Procedure

Child participants and their parents were mailed separate packets of questionnaires (with separate self-addressed return envelopes) so that children could keep their responses private from their parents (and this was suggested in instructions to both children and parents). Written consent was obtained from parents and written assent from children. Parents were asked to give the teacher version of the Social Skills Improvement System to their child's teacher in the humanities and/or math. Packets also included instructions on how to access the online EF tasks. Children were instructed to complete the tasks when they were alone and free from distractions.

Questionnaires

Participants and their parents and teachers completed a battery of questionnaires assessing social skills, sensation seeking, and academic performance.

Social skills and academics. *The Social Skills Improvement System* (SSIS; Gresham and Elliot, 2008) assessed children's social functioning in everyday life and was completed by parents, children, and teachers separately (three versions created by the developers). The form contains 75–85 questions, depending on the informant. Questions (e.g., "Takes responsibility for part of a group activity.") are rated on a 4-point scale (never, sometimes, often, almost always). Subscales for social skills (46 items), problem behaviors (30 items), and academic competence (teacher form only) are included. For the academic competence scale, teachers rated how children ranked among their peers on a 5-point scale (ranging from lowest 10% to highest 10%) for 7 items that queried specific academic skills, motivation, and intellectual ability. Internal consistency alphas for each subscale of this form are 0.93–0.96 (Gresham and Elliot, 2008). For the current study, we included child and teacher reports in our analyses. Higher

scores on the social skills and academic performance scales indicate better performance. We do not report on problem behaviors, because teachers reported very low levels of problem behaviors in this sample (mean 8.7, whereas rating "sometimes" for each item would score 30).

Sensation seeking. Children completed the *Sensation Seeking Scale for Children* (SSSC; Russo et al., 1993). In this 26-item form, children are asked to choose between two alternatives, e.g., "I don't do anything I might get in trouble for" vs. "I like to do new and exciting things, even if I think I might get in trouble for doing them." The form has subscales for thrill and adventure seeking, drug and alcohol seeking, and social disinhibition. We collected data on the full questionnaire, but used only the thrill and adventure subscale (12 items) for analyses because children in this sample reported little sensation seeking in the other two categories. The internal consistency alpha reported for this subscale is 0.81 (Russo et al., 1993). Scores for this subscale were summed, and higher scores reflect higher levels of sensation seeking.

Online EF games. Participants completed online, computerized versions of the NIH Toolbox DCCS and Flanker tasks (Zelazo et al., 2013). The Flanker task follows a similar format as the ANT task used at Time 1 except that the cue was always a central star and stimuli were arrows instead of fish. Participants completed 4 practice trials and 20 test trials. The DCCS was the same format as at Time 1: participants again sorted stimuli by shape (dominant) or color (non-dominant). They completed eight practice trials (four sorting by shape and four sorting by color) and 30 test trials, which included 80% dominant cues and 20% non-dominant cues. A combined score that took into account accuracy and reaction times was calculated for each of the two tasks (theoretical range 0–10).

RESULTS

We first examined effects of age and gender on variables of interest performance on EF tasks, P300, SPN, sensation seeking, academic performance, and social skills. We then examined stability of cool EF performance from age 8–12, followed by concurrent and longitudinal links between cool EF and adaptive behavior. Finally, we examined whether performance on the HDT task and neural correlates predicted later outcomes.

PRELIMINARY ANALYSES

We examined descriptive statistics for the Flanker and DCCS at age 8 and 12 (**Table 1**). At age 8, reaction times for incongruent/non-dominant trials were negatively correlated with accuracy, indicating that children slowed down to perform well on the tasks at this age. Therefore, we used percent accuracy on these more difficult trials as predictors of future EF and adaptive behavior. For age 12, accuracy scores reached ceiling, so we used a composite of accuracy and RT using the NIH toolbox algorithm.

We examined whether verbal ability (assessed at age 8 using the *PPVT-4*) was correlated with EF performance and questionnaire scores. Verbal ability was not significantly correlated with any variables of interest (r's $= -0.002$ to 0.25). However, it was marginally correlated with academic performance ($r = 0.25$, $p < 0.09$), so we controlled for verbal ability when examining correlations with academic performance.

Table 1 | Descriptive statistics for executive function variables.

	N	Theoretical range	Mean	Std. Dev
Age 8: Flanker incongruent accuracy	75	0–1	0.95	0.07
Age 8: DCCS non-dominant accuracy	77	0–1	0.83	0.11
Age 12: Flanker score	67	0–10	7.79	0.45
Age 12: DCCS score	67	0–10	7.28	0.63

Note: Only children who participated at both time points are included.

In addition, we examined gender differences for each variable. Girls obtained significantly higher scores on the DCCS [$F_{(1, 65)} = 3.06$, $p < 0.01$] and self-reported better social skills [$F_{(1, 75)} = 6.5$, $p < 0.02$] at age 12. No other significant gender differences were found.

STABILITY OF EF FROM AGE 8 TO 12

To examine the stability of EF, we included only the 67 children who completed at least one EF task at both time points. At age 8, Flanker and DCCS performance were measured using accuracy on incongruent and non-dominant (color) trials, respectively. Accuracies for incongruent/non-dominant trials on the two tasks were not significantly correlated at this age [$r_{(66)} = 0.20$, $p = 0.1$], although this is likely due to a ceiling effect on Flanker task accuracy at age 8. Reaction times were significantly correlated across the two tasks [$r_{(66)} = 0.29$, $p < 0.02$]. RTs were positively correlated with accuracy [Flanker: $r_{(66)} = 0.27$, $p < 0.03$; DCCS: $r_{(66)} = 0.41$, $p = 0.001$], indicating that children at this age slowed down to achieve better performance, whereas at later ages, faster RTs indicate greater efficiency.

At age 12, accuracy on these tasks reached ceiling levels, so performance was measured using an algorithm from the NIH toolbox (Zelazo et al., 2013) that combined accuracy and reaction time (in which participants receive a higher score for responding quickly after full accuracy is reached). This algorithm computes a score from 0 to 10 (sample range = 5–8.67). At age 12, performance on the DCCS and Flanker were uncorrelated, $r_{(67)} = 0.015$. As shown in **Table 2**, age 8 Flanker accuracy predicts age 12 Flanker, but not DCCS performance, while age 8 DCCS accuracy predicts age 12 DCCS, but not Flanker performance. RTs at age 8 were not significantly correlated with performance at age 12 (ps > 0.08). These results indicate longitudinal stability within but not across each EF task.

CONCURRENT LINKS BETWEEN EF AND ADAPTIVE BEHAVIOR AT AGE 12

We examined links between our EF measures and self- and teacher-reported social skills, thrill/adventure seeking (self-report only), and academic competence (teacher report only) at age 12. DCCS performance was positively correlated with self-reported social skills, $r_{(64)} = 0.29$, $p = 0.02$, but negatively correlated with academic competence $r_{(42)} = -0.32$, $p = 0.04$ (**Table 2**). Because

Table 2 | Correlations between age 8 and age 12 variables.

	Age 8 Flanker (ANT)	Age 8 HDT P3 Effect	Age 8 HDT SPN Effect	Age 8 HDT Losses	Age 12 DCCS	Age 12 Flanker	Age 12 TAS	Age 12 Child Soc. Skills	Age 12 Teacher Soc. Skills	Age 12 Academic Perf. (PPVT Control)
Age 8 DCCS	0.216 $n = 75$	0.108 $n = 74$	−0.030 $n = 75$	−0.089 $n = 76$	0.278* $n = 66$	−0.123 $n = 66$	−0.210 $n = 75$	0.016 $n = 75$	−0.133 $n = 66$	−0.218(0.069) $n = 62(52)$
Age 8 Flanker (ANT)	–	0.005 $n = 72$	−0.026 $n = 73$	−0.21 $n = 74$	−0.059 $n = 64$	0.381** $n = 64$	−0.085 $n = 75$	0.027 $n = 75$	0.139 $n = 64$	0.35**(0.35)** $n = 62(52)$
Age 8 HDT P3 Effect		–	0.144 $n = 74$	−0.111 $n = 74$	0.166 $n = 63$	−0.037 $n = 63$	−0.250* $n = 73$	−0.123 $n = 73$	0.102 $n = 65$	0.061(0.053) $n = 62(44)$
Age 8 HDT SPN Effect			–	−0.074 $n = 75$	−0.080 $n = 63$	−0.007 $n = 63$	0.091 $n = 73$	−0.148 $n = 73$	0.139 $n = 65$	0.281*(0.31)* $n = 62(45)$
Age 8 HDT Losses				–	−0.072 $n = 65$	0.077 $n = 65$	0.104 $n = 75$	−0.001 $n = 75$	0.065 $n = 66$	−0.17(−0.18) $n = 62(44)$
Age 12 DCCS					–	0.015 $n = 67$	−0.071 $n = 63$	0.292* $n = 64$	−0.052 $n = 54$	−0.32*(−0.33)* $n = 51(48)$
Age 12 Flanker						–	0.072 $n = 63$	−0.086 $n = 63$	0.032 $n = 54$	−0.106(−0.142) $n = 51(48)$
Age 12 TAS							–	−0.174 $n = 77$	−0.03 $n = 62$	0.151(0.144) $n = 58(55)$
Age 12 Child Soc. Skills								–	0.348** $n = 62$	0.191(0.161) $n = 58(55)$
Age 12 Teacher Soc. Skills									–	0.61**(0.58)** $n = 62(55)$

*$p < 0.05$, **$p < 0.01$.

Correlations are computed based on children who participated at both age 8 and age 12. Correlations involving academic competence are presented both raw and partial (controlling for PPVT score). TAS, Thrill/Adventure Seeking.

there was an effect of gender on DCCS performance at age 12, we performed a partial correlation controlling for gender, which did not affect the magnitude of these correlations. Flanker performance was not significantly correlated with any outcome variables.

LINKS BETWEEN AGE 8 EF AND AGE 12 ADAPTIVE BEHAVIOR

We tested the degree to which accuracy on EF tasks at age 8 predicts adaptive behavior at age 12. Flanker incongruent trial accuracy significantly predicted teacher-rated academic competence, but no other outcome variables. This correlation remained significant when controlling for verbal ability and age 12 Flanker performance (see **Table 2**). DCCS non-dominant trial accuracy did not predict any outcome variables.

LONGITUDINAL PREDICTIONS OF AFFECTIVE DECISION-MAKING

Next, data were analyzed to assess the degree to which performance on and neural correlates of a risky decision-making task at age 8 (Hungry Donkey) predicted individual differences in variables of interest at age 12. To measure performance, we examined the total number of trials in which a net loss was incurred. Two neural correlates were of interest, stemming from our previous findings (Carlson et al., 2009): (i) the magnitude of the post-stimulus P300 and (ii) the pre-stimulus/anticipatory SPN components in response to reward and loss trials.

The P300 was of significantly larger magnitude in response to loss vs. reward trials, $[F_{(1, 78)} = 31.2, p < 0.001]$ and the SPN was significantly larger (more negative-going) after high-frequency loss door selections than low-frequency loss door selections $[F_{(1, 78)} = 6.51, p < 0.02]$. However, our primary question was whether individual differences in the magnitude of P300 to loss trials and SPN to high-frequency loss door selections predicted adaptive outcomes at age 12. Individual differences in P300 effect (magnitude to loss-minus-reward trials) was negatively correlated with self-reported thrill/adventure seeking. In other words, the larger the neural response to punishment (vs. reward) outcomes at age 8, the less likely participants were to report an interest in thrill/adventure 4 years later in pre-adolescence. The P300 effect was not significantly correlated with other outcome variables.

In addition, individual differences in SPN magnitude to low-minus-high-frequency loss doors, in which more positive values reflect pre-outcome anticipation of loss, significantly predicted

academic competence at age 12, even when controlling for verbal ability. This finding suggests that a neural correlate of risk-aversion at age 8 was related to higher academic competence at age 12, but not to other outcome variables. The P300 and SPN components did not predict EF performance at age 12, and the loss count on the HDT task did not predict any outcomes at age 12 (see **Table 2** for summary).

DISCUSSION

The goals of this research were to examine the stability of EF and the extent to which individual differences in neural responses to affective cues serve as a biological marker for EF and adaptive behaviors in a sample followed longitudinally from age 8 to 12. Two broad findings emerged. We found that (i) cool aspects of EF showed modest stability from middle childhood to early adolescence and that (ii) certain aspects of childhood cool EF and neural sensitivity to reward and punishment predicted some individual differences in sensation seeking and adaptive behaviors in children entering adolescence.

STABILITY OF COOL EF AND LINKS TO ADAPTIVE BEHAVIOR

To our knowledge, this is the first study to demonstrate longitudinal stability of performance on specific EF tasks in this age range from middle childhood to adolescence. Our findings add to a growing body of evidence suggesting that individual differences in EF remain stable beyond the preschool period (Eigsti et al., 2006; Casey et al., 2011). In our sample, the inhibition (ANT/Flanker) and shifting/updating (DCCS) aspects of EF showed stability across the age range tested: Individual differences in ANT performance at age 8 predicted Flanker performance at age 12, and DCCS performance at age 8 predicted DCCS performance at age 12. However, performance on these two tasks were uncorrelated with each other in this age range. Given that different dimensions of cool EF tend to be strongly correlated in the preschool years (e.g., Wiebe et al., 2011), our findings suggest that dimensions of cool EF may become more differentiated over time, which is compatible with prior cross-sectional work reporting separability of working memory updating and inhibitory control beginning around age 9–10 (Shing et al., 2010). Taken together, these results support an idea of increasing specialization of circuits within the PFC for specific cognitive functions across development (e.g., Zelazo and Carlson, 2012).

A noteworthy aspect of our longitudinal design is that we documented that inhibitory control (ANT) at age 8 predicted teacher-rated academic competence at age 12. This finding fits with other work linking EF and later academic achievement in a variety of age groups (Blair and Razza, 2007; Best et al., 2011). Links between early EF and later academic performance make sense, given that the abilities to inhibit prepotent responses and to ignore distractions are necessary to develop the self-control necessary to be attentive in class and to study or do homework instead of engaging in other activities. The fact that this finding was independent of verbal ability (PPVT) lends further support to the emerging belief that EF matters for later academic achievement over and above, and perhaps more than IQ (for review, see Duckworth and Carlson, 2013).

NEURAL SENSITIVITY TO PUNISHMENT vs. REWARD PREDICTS LATER BEHAVIOR

Neural responses during the child-friendly gambling task (a hot EF task) at age 8 predicted a variety of adaptive behaviors at age 12. Greater sensitivity to loss trials and high-frequency loss doors, as indexed by the magnitude of P300 and SPN difference scores, predicted lower propensity for thrill/adventure seeking and better academic outcomes, respectively. Interestingly, avoidance of losses during the task did not correlate with later outcomes, suggesting that our ERP measures were more sensitive than behavior. Although children showed sensitivity to rewards and punishments at a neural level, they may not have been fully able to translate that sensitivity into improved performance during the course of the task. However, our results suggest that greater sensitivity to punishment vs. reward may play an important role in the development of children's trajectories toward more cautious/conscientious vs. higher risk-taking behavioral patterns.

Children who had shown attenuated P300 amplitudes after loss trials relative to reward trials reported more desire to engage in risky behaviors in early adolescence. This finding fits well with results linking reduced P300 amplitude to risky behaviors such as alcohol use in adolescence (Carlson et al., 1999; McGue et al., 2001). Although ours was a low-risk sample that reported low levels of externalizing behaviors in general, these findings add to evidence that the P300 could be a psychophysiological marker linked to propensity for risk-taking and/or sensation seeking. These findings also suggest that children who devoted more attentional resources to loss trials (reflected in *higher* P300 amplitudes) may tend toward more risk-averse behavioral trajectories.

While the post-feedback P300 at 8 years old was associated with adolescent thrill/adventure seeking, it is interesting that pre-feedback SPN at 8 years old predicted academic outcomes in adolescence. Specifically, we found that greater magnitude of SPN responses to high-frequency loss doors predicted greater academic success. The SPN is believed to index a "somatic marker" (Damasio, 1996) involving cortical and subcortical activity related to the expectation for relevant positive or negative feedback (Brunia et al., 2011). For example, it is enhanced under conditions in which outcomes are linked to actions vs. occurring at random (Masaki et al., 2010), suggesting that a sense of control during a task is necessary to elicit the SPN. In our study, SPN responses tended to be larger just before high-probability loss outcomes than low-probability loss outcomes. On the surface, our findings might seem to contradict those of Stavropoulos and Carver (2013), who found that anticipation of more socially rewarding vs. less rewarding feedback elicited a larger SPN in children. However, both findings make sense if SPN is a marker of the salience or motivational relevance of stimuli. While the former study did not involve punishment, ours did, so children were likely motivated primarily to avoid losses. In the current study, children who showed a larger SPN in the moment just after making a risky selection and just before a loss outcome was revealed might be more sensitive to the fact that they made a non-optimal response and therefore expect to receive negative feedback. In other words, they had a "feeling" or intuition detectable at a neural level that they were about to suffer a loss on the next trial. The present study suggests that perhaps they felt they had agency to

avoid future losses and were better able to learn from mistakes in general, which may in turn facilitate learning from pedagogical instruction and success in school.

This interpretation about the SPN is speculative, in part because the interpretation and study of SPN is just beginning to be applied in children (e.g., Stavropoulos and Carver, 2013). However, a cross-sectional study recently showed that activity in insular cortex, believed to be a primary generator of the SPN, increased between age 5 and adulthood during a gambling task, corresponding to an age-related increase in risk aversion (Paulsen et al., 2011). In addition, emerging evidence links the SPN to the dopaminergic learning systems believed to underlie the error-related negativity component (ERN) in adults (Moris et al., 2013). The ERN varies in magnitude according to the difference between expected and received feedback and has been better characterized developmentally than the SPN. Examining both of these components in development and linking them to laboratory and real-world behavior would yield rich information about how neural sensitivity to reward and punishment relates to learning and adaptation.

LIMITATIONS AND FUTURE DIRECTIONS

This research has limitations and suggests future directions. Although we found that cool EF and ERP components during a hot EF task predicted some individual differences 4 years later, we found no evidence to support other predicted longitudinal relations. Behavioral performance on the HDT did not predict later outcomes, suggesting that on this task, ERP components were a more sensitive measure of individual differences in sensitivity to reward and punishment. However, these neural measures did not predict later social skills, which were related (for self-report) only to concurrent DCCS performance. Surprisingly, at the same time, age 12 DCCS performance was *negatively* correlated with teacher-rated academic performance. Self-rated social skills and teacher-rated academic competence were not significantly correlated, which is not surprising at this age in a low-risk sample. Possibly, children in our sample who are more successful academically were more conscientious in general, and this negatively affected their score on the DCCS, where slowing down in order to be accurate would result in a cost. We also found no evidence that cool EF at either time point predicted thrill/adventure seeking, but this null finding fits with developmental literature showing evidence of a dissociation between impulsivity and sensation seeking (Steinberg et al., 2008).

Another limitations is that our sample was relatively homogeneous in terms of ethnic and socioeconomic background. Children were generally low-risk and not (yet) endorsing many of the substance use and social risk-taking behaviors on the SSSC. We are continuing to follow this cohort through adolescence when some of these items will become more sensitive. Nonetheless, there were sufficient individual differences in academic competence and thrill/adventure seeking (e.g., enjoyment of riding one's bike fast down a steep hill) to detect longitudinal predictions from 4 years earlier. As well, we did not use the same versions of the EF tasks at age 8 and 12 in this longitudinal sample because the abbreviated NIH Toolbox versions were not available at Time 1 and developed in the interim. Nevertheless, the within-task stability of the Flanker and DCCS was significant. Finally, given our relatively small sample size, correction for multiple comparisons would have reduced some findings to non-significance, but we note that we were selective in our comparisons and had a priori hypotheses regarding each of them.

Despite these limitations, this is the first longitudinal study, to our knowledge, to examine the development of both hot and cool EF and their relations to adaptive behavior between middle childhood and pre-adolescence. We found that individual differences in performance on EF tasks were stable across this age range and that certain aspects of cool and hot EF predicted individual differences in thrill/adventure seeking and academic outcomes at age 12.

These novel findings generate many potential directions for future research, especially regarding adolescent brain development and the prediction of individual differences in adaptive behavior. Hot and cool EF appear to interact in complex ways that change across development and these aspects of EF may relate to behavior differently in childhood, adolescence, and adulthood. For example, hot EF may become increasingly important relative to cool EF in adolescence, when individuals begin to take greater control of their environment and make more decisions for themselves. In addition, adolescents show more sensitivity to social facilitation from peers than children or adults in the context of a risky decision-making task (Gardner and Steinberg, 2005). Therefore, we might expect peers to play a larger role in either facilitating or hindering adaptive behavior in adolescence than in other age groups. With this in mind, both social understanding and EF may be key to the successful navigation through adolescence. Future research could also more deeply explore the relations between EF and risk aversion, as opposed to risk-taking. Such work could have implications for anxiety disorders, which may be linked to maladaptive levels of risk aversion (Robin and Martin, 2010) and are especially prevalent in the teenage years. Exploration of the role of hot EF in development, particularly at a neural level, is a new area that holds great promise for deepening our understanding of human brain-behavior relations, and we expect that future studies will yield information with high applicability for developmental theory, educational practice, and clinical science.

AUTHOR CONTRIBUTIONS

Madeline B. Harms wrote the majority of the manuscript and contributed to the study design and data collection at Time 2. Vivian Zayas contributed to the study design and data collection at Time 1. This research was conducted in the lab of Stephanie M. Carlson and she contributed to the study design and data collection at both time points. All four authors contributed to analysis and interpretation of data and revising the manuscript, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

The authors wish to thank Catherine Schaefer for extensive work in data collection, scoring, and entry, as well as Jacob Anderson and Max Shinn for assisting in design of scoring algorithms. This work was funded in part from a grant from the University of

Washington Institute for Learning and Brain Sciences, a Royalty Research Fund award from the University of Washington, and R01HD051495 to Stephanie M. Carlson.

REFERENCES

Albert, D., and Steinberg, L. (2011). Judgment and decision making in adolescence. *J. Res. Adolesc.* 21, 211–224. doi: 10.1111/j.1532-7795.2010.00724.x

Bechara, A. (2004). The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. *Brain Cogn.* 55, 30–40. doi: 10.1016/j.bandc.2003.04.001

Best, J. R., Miller, P. H., and Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learn Individ. Differ.* 21, 327–336. doi: 10.1016/j.lindif.2011.01.007

Beyth-Marom, R., Austin, L., Fischhoff, B., Palmgren, C., and Jacobs-Quadrel, M. (1993). Perceived consequences of risky behaviors: adults and adolescents. *Dev. Psychol.* 29, 549–563. doi: 10.1037/0012-1649.29.3.549

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Bocker, K. B. E., Brunia, C. H. M., and Van den Berg-Lenssen, M. M. C. (1994). A spatiotemporal dipole model of the Stimulus Preceding Negativity (SPN) prior to feedback stimuli. *Brain Topogr.* 7, 71–88. doi: 10.1007/BF01184839

Brunia, C. H. M., Hackley, S. A., van Boxtel, G. J. M., Kotani, Y., and Ohgami, Y. (2011). Waiting to perceive: reward or punishment? *Clin. Neurophysiol.* 122, 858–868. doi: 10.1016/j.clinph.2010.12.039

Bunge, S. A., and Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Curr. Dir. Psychol. Sci.* 15, 118–121. doi: 10.1111/j.0963-7214.2006.00419.x

Carlson, S. M., Mandell, D. J., and Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Dev. Psychol.* 40:1105. doi: 10.1037/0012-1649.40.6.1105

Carlson, S. M., and Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333

Carlson, S. M., White, R. E., and Davis-Unger, A. C. (2014). Evidence for a relation between executive function and pretense representation in preschool children. *Cogn. Dev.* 29, 1–16. doi: 10.1016/j.cogdev.2013.09.001

Carlson, S. M., Zayas, V., and Guthormsen, A. (2009). Neural correlates of decision making on a gambling task. *Child Dev.* 80, 1076–1096. doi: 10.1111/j.1467-8624.2009.01318.x

Carlson, S. M., Zelazo, P. D., and Faja, S. (2013). "Executive function," in *The Oxford Handbook of Developmental Psychology, Vol. 1: Body and Mind*, ed P. D. Zelazo (New York, NY: Oxford University Press), 706–743.

Carlson, S. R., Katsanis, J., Iacono, W. G., and Mertz, A. K. (1999). Substance abuse and externalizing psychopathology in adolescent boys with small, average, or large P300 event-related potential amplitude. *Psychophysiology* 36, 249–255. doi: 10.1111/1469-8986.3650583

Casey, B. J., Giedd, J. N., and Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biol. Psychol.* 54, 241–257. doi: 10.1016/S0301-0511(00)00058-2

Casey, B. J., Somerville, L. H., Gotlib, I. A., Ayduk, O., Franklin, N. T., Askren, M. K. et al. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14998–15003. doi: 10.1073/pnas.1108561108

Cauffman, E., Shulman, E. P., Claus, E., Banich, M. T., Steinberg, L., Graham, S., et al. (2010). Decision making as indexed by performance on the Iowa Gambling Task. *Dev. Psychol.* 46, 193–207. doi: 10.1037/a0016128

Crone, E. A., and van der Molen, M. W. (2004). Developmental changes in real-life decision-making: performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Dev. Neuropsychol.* 25, 251–279. doi: 10.1207/s15326942dn2503_2

Crone, E. A., and van der Molen, M. W. (2007). Development of decision making in school-aged children and adolescents: evidence from heart rate and skin conductance analysis. *Child Dev.* 78, 1288–1301. doi: 10.1111/j.1467-8624.2007.01066.x

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 351, 1413–1420. doi: 10.1098/rstb.1996.0125

Duckworth, A. L., and Carlson, S. M. (2013). "Self-regulation and school success," in *Self-Regulation and Autonomy: Social and Developmental Dimensions of Human Conduct*, eds B. W. Sokol, F. M. E. Grouzet, and U. Müller (New York, NY: Cambridge University Press), 208–230. doi: 10.1017/CBO9781139152198.015

Dunn, D. M., and Dunn, L. M. (2007). *Peabody Picture Vocabulary Test, 4th Edn. manual*. Minneapolis, MN: NCS Pearson, Inc.

Eigsti, I., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M. B., et al. (2006). Predictive cognitive control from preschool to late adolescence and adulthood. *Psychol. Sci.* 17, 478–484. doi: 10.1111/j.1467-9280.2006.01732.x

Eluvathingal, T., Hasan, K., Kramer, L., Fletcher, J., and Ewing-Cobbs, L. (2007). Quantitative diffusion tensor tractography of association and projection fibers in normally developing children and adolescents. *Cereb. Cortex* 17, 2760–2768. doi: 10.1093/cercor/bhm003

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non-search task. *Percept. Psychophysiol.* 16, 143–149. doi: 10.3758/BF03203267

Ernst, M., Romeo, R. D., and Andersen, S. L. (2009). Neurobiology of the development of motivated behaviors in adolescence: a window into a neural systems model. *Pharmacol. Biochem. Behav.* 93, 199–211. doi: 10.1016/j.pbb.2008.12.013

Finer, L., and Henshaw, S. (2006). Disparities in rates of unintended pregnancy in the United States, 1994 and 2001. *Perspect. Sex. Reprod. Health* 38, 90–96. doi: 10.1363/3809006

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., et al. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk taking in adolescents. *J. Neurosci.* 26, 6885–6892. doi: 10.1523/JNEUROSCI.1062-06.2006

Gardner, M., and Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Dev. Psychol.* 41, 625–635. doi: 10.1037/0012-1649.41.4.625

Gresham, F. M., and Elliot, S. N. (2008). *Social Skills Improvement System: Rating Scales and Manual.* Minneapolis, MN: Pearson Assessments.

Groen, Y., Wijers, A. A., Mulder, L. J. M., Minderaa, R. B., and Althaus, M. (2007). Physiological correlates of learning by performance feedback in children: a study of EEG event-related potentials and evoked heart rate. *Biol. Psychol.* 76, 174–187. doi: 10.1016/j.biopsycho.2007.07.006

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., and Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: age-related changes and individual differences. *Dev. Neuropsychol.* 28, 617–644. doi: 10.1207/s15326942dn2802_4

Hooper, C. J., Luciana, M., Conklin, H. M., and Yarger, R. S. (2004). Adolescents' performance on the Iowa Gambling Task: implications for the development of decision-making and ventromedial prefrontal cortex. *Dev. Psychol.* 40, 1148–1158. doi: 10.1037/0012-1649.40.6.1148

Hughes, C., and Ensor, R. (2007). Executive function and theory of mind: predictive relations from ages 2 to 4. *Dev. Psychol.* 43, 1447–1459. doi: 10.1037/0012-1649.43.6.1447

Huizenga, H. M., Crone, E. A., and Jansen, B. J. (2007). Decision-making in healthy children, adolescents, and adults explained by the use of increasingly complex proportional reasoning rules. *Dev. Sci.* 10, 814–825. doi: 10.1111/j.1467-7687.2007.00621.x

Masaki, H., Yamazaki, K., and Hackley, S. H. (2010). Stimulus-preceding negativity is modulated by action-outcome contingency. *Neuroreport* 21, 277–281. doi: 10.1097/WNR.0b013e3283360bc3

McGue, M., Iacono, W. G., Legrand, L. N., Malone, S., and Elkins, I. (2001). Origins and consequences of age at first drink. I. Associations with substance-use disorders, disinhibitory behavior and psychopathology, and P3 amplitude. *Alcohol. Clin. Exp. Res.* 25, 1156–1165. doi: 10.1111/j.1530-0277.2001.tb02330.x

Mischel, W., Shoda, Y., and Rodriguez, M. L. (1989). Delay of gratification in children. *Science* 244, 933–938. doi: 10.1126/science.2658056

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their

contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moris, J., Luque, D., and Rodriguez-Fornells, A. (2013). Learning-induced modulations of the stimulus-preceding negativity. *Psychophysiology* 50, 931–939. doi: 10.1111/psyp.12073

Paulsen, D. J., Carter, R. M., Platt, M. L., Huettel, S. A., and Brannon, E. M. (2011). Neurocognitive development of risk aversion from early childhood to adulthood. *Front. Hum. Neurosci.* 5:178. doi: 10.3389/fnhum.2011.00178

Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends Cogn. Sci.* 9, 60–68. doi: 10.1016/j.tics.2004.12.008

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., and Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *J. Exp. Child Psychol.* 108, 621–637. doi: 10.1016/j.jecp.2010.09.008

Reyna, V. F., and Zayas, V. (eds.) (2014).*The Neuroscience of Risky Decision.* Washington, DC: American Psychological Association. doi: 10.1037/14322-000

Robin, A. L., and Martin, P. P. (2010). Neural systems underlying approach and avoidance in anxiety disorders. *Dialogues Clin. Neurosci.* 12, 517–531.

Rueda, M. R., Fan, J., McCandlis, B. D., Halparin, J. D., Gruber, D. B., Pappert Lercari, L., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Russo, M. F., Stokes, G. S., Lahey, B. B., Christ, M. A. G., McBurnett, K., Loeber, R., et al. (1993). A sensation seeking scale for children: refinement and psychometric development. *J. Psychopathol. Behav. Assess.* 15, 69–86. doi: 10.1007/BF00960609

Seeyave, D. M., Coleman, S., Appugliese, D., Corwyn, R. F., Bradley, R. H., Davidson, N. S., et al. (2009). Ability to delay gratification at age 4 years and risk of overweight at age 11 years. *Arch. Pediatr. Adolesc. Med.* 163, 303–308. doi: 10.1001/archpediatrics.2009.12

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S., and Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Dev. Neuropsychol.* 35, 679–697. doi: 10.1080/87565641.2010.508546

Smith, D. G., Xiao, L., and Bechara, A. (2012). Decision making in children and adolescents: impaired Iowa Gambling Task performance in early adolescence. *Dev. Psychol.* 48, 1180–1187. doi: 10.1037/a0026342

Stavropoulos, K. K. M., and Carver, L. J. (2013). Reward sensitivity to faces versus objects in children: an ERP study. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst149. [Epub ahead of print].

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends Cogn. Sci.* 9, 69–74. doi: 10.1016/j.tics.2004.12.005

Steinberg, L. (2010). A dual systems model of adolescent risk-taking. *Dev. Psychobiol.* 52, 216–224. doi: 10.1002/dev.20445

Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., and Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. *Dev. Psychol.* 44, 1764–1778. doi: 10.1037/a0012955

Substance Abuse and Mental Health Services Administration. (2007). *Results from the 2006 National Survey on Drug Use and Health: National findings*. Rockville, MD: Substance Abuse and Mental Health Services Administration

Tucker, D. M. (1993). Spatial sampling of electrical fields: the geodesic sensor net. *Electroencephalogr. Clin. Neurophysiol.* 87, 154–163. doi: 10.1016/0013-4694(93)90121-B

Vaish, A., Grossman, T., and Woodward, A. (2008). Not all emotions are created equal: the negativity bias in social-emotional development. *Psychol. Bull.* 134, 383–403. doi: 10.1037/0033-2909.134.3.383

Van Leijenhorst, L., Crone, E. A., and Bunge, S. A. (2006). Neural correlates of developmental differences in risk estimation and feedback processing. *Neuropsychologia* 44, 2158–2170. doi: 10.1016/j.neuropsychologia.2006.02.002

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., and Espy, K. A. (2011). The structure of executive function in 3-year-olds. *J. Exp. Child Psychol.* 108, 436–452. doi: 10.1016/j.jecp.2010.08.008

Zayas, V., Mischel, W., and Pandey, G. (2014). "Mind, brain, and delay of gratification," in *The Neuroscience of Risky Decision*, eds V. F. Reyna and V. Zayas (Washington, DC: American Psychological Association). doi: 10.1037/14322-007

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., and Weintraub, S. (2013). NIH toolbox Cognition Battery (CB): measuring executive function and attention. *Monogr. Soc. Res. Child Dev.* 78, 16–33. doi: 10.1111/mono.12032

Zelazo, P. D., and Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence. *Child Dev. Perspect.* 6, 354–360. doi: 10.1111/j.1750-8606.2012.00246.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 January 2014; accepted: 31 March 2014; published online: 22 April 2014.
Citation: Harms MB, Zayas V, Meltzoff AN and Carlson SM (2014) Stability of executive function and predictions to adaptive behavior from middle childhood to pre-adolescence. Front. Psychol. ***5****:331. doi: 10.3389/fpsyg.2014.00331*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Harms, Zayas, Meltzoff and Carlson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 01 July 2014
doi: 10.3389/fpsyg.2014.00496

Trail making test performance in youth varies as a function of anatomical coupling between the prefrontal cortex and distributed cortical regions

Nancy Raitano Lee[1]*, Gregory L. Wallace[2], Armin Raznahan[1], Liv S. Clasen[1] and Jay N. Giedd[1]

[1] *Child Psychiatry Branch, Intramural Research Program, National Institute of Mental Health, NIH, Bethesda, MD, USA*
[2] *Department of Speech and Hearing Sciences, George Washington University, Washington, DC, USA*

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Joseph M. Orr, University of Colorado at Boulder, USA
Sandra Wiebe, University of Alberta, Canada

****Correspondence:***
Nancy Raitano Lee, Child Psychiatry Branch, National Institute of Mental Health, NIH, Bldg., 10, Room 4C110, 10 Center Drive, MSC 1367, Bethesda, MD 20892, USA
e-mail: lnancy@mail.nih.gov

While researchers have gained a richer understanding of the neural correlates of executive function in adulthood, much less is known about how these abilities are represented in the developing brain and what structural brain networks underlie them. Thus, the current study examined how individual differences in executive function, as measured by the Trail Making Test (TMT), relate to structural covariance in the pediatric brain. The sample included 146 unrelated, typically developing youth (80 females), ages 9–14 years, who completed a structural MRI scan of the brain and the Halstead-Reitan TMT (intermediate form). TMT scores used to index executive function included those that evaluated set-shifting ability: Trails B time (number-letter sequencing) and the difference in time between Trails B and A (number sequencing only). Anatomical coupling was measured by examining correlations between mean cortical thickness (MCT) across the entire cortical ribbon and individual vertex thickness measured at ∼81,000 vertices. To examine how TMT scores related to anatomical coupling strength, linear regression was utilized and the interaction between age-normed TMT scores and both age and sex-normed MCT was used to predict vertex thickness. Results revealed that stronger Trails B scores were associated with greater anatomical coupling between a large swath of prefrontal cortex and the rest of cortex. For the difference between Trails B and A, a network of regions in the frontal, temporal, and parietal lobes was found to be more tightly coupled with the rest of cortex in stronger performers. This study is the first to highlight the importance of structural covariance in in the prediction of individual differences in executive function skills in youth. Thus, it adds to the growing literature on the neural correlates of childhood executive functions and identifies neuroanatomic coupling as a biological substrate that may contribute to executive function and dysfunction in childhood.

Keywords: executive function, anatomical covariance, cortical thickness, magnetic resonance imaging, Trail Making Test, brain, child, adolescent

INTRODUCTION

For over 150 years, scientists studying cognition have noted the important role of the frontal lobes in the regulation of behavior and cognition (see Braver and Ruge, 2006 for a review). While early investigations focused mainly on adult clinical populations, more recent research has described the protracted development of both executive functions and the frontal lobes within the context of normative development.

Executive function is an umbrella term referring to a collection of skills (such as working memory, planning, inhibition, and cognitive flexibility) that are thought to be essential for solving unfamiliar problems and coping with changing demands in one's environment (Lezak et al., 2004). Normative studies indicate that executive function skills develop across childhood and into early young adulthood, with different skills reaching "mature" adult levels at different points in development.

Studies of the protracted nature of the development of executive functions within the context of typical development span several decades. Starting with the early work of Welsh and Pennington (1988) and continuing to more recent investigations (Hooper et al., 2004; Luciana et al., 2005; Huizinga et al., 2006; Conklin et al., 2007), a large corpus of data now exists documenting that youth continue to make gains in performance on several different executive function tasks into the mid to late teens (Luna et al., 2004; Luciana et al., 2005).

Complementing these behavioral studies, morphometric studies of the developing brain using structural magnetic resonance imaging have suggested that the prefrontal cortex, thought to be central to the executive functions, is among the latest maturing regions of the brain (e.g., Gogtay et al., 2004). Its protracted development contrasts with the relatively early development of brain regions thought to contribute to more basic sensory and motor functions, such as the somatosensory cortex.

Thus, both behavioral and anatomical data suggest that childhood and adolescence are times in which studies of the anatomical correlates of executive abilities may be most informative

in augmenting our understanding of how higher-level cognitive abilities develop typically and atypically. With regard to atypical executive development, most, if not all, developmental disorders are characterized by executive deficits. Examples include attention-deficit/hyperactivity disorder, conduct disorder, autism spectrum disorders, and intellectual disability, to name a few (for reviews, see Pennington and Ozonoff, 1996; Zelazo and Muller, 2002). Furthermore, many psychiatric disorders that develop in late adolescence or early adulthood, such as schizophrenia and depression, are characterized by executive deficits (Orellana and Slachevsky, 2013; Snyder, 2013).

Understanding the neuroanatomical correlates of executive abilities within the context of typical development may inform research seeking to identify mechanisms that contribute to the atypical development of executive functioning in childhood or in disorders that first manifest in adolescence/early adulthood. In the current investigation, we focus on the neuroanatomical correlates of a commonly-used measure of executive function, the Trail Making Test (TMT), in a sample of typically-developing youth, ages 9–14 years. The TMT, like many neuropsychological assessment tools, was first developed for adult populations. The original task, called the Pathways Test, was included in the Army Individual Test of General Ability in the 1940s (Partington and Leiter, 1949). The TMT is probably best known as being a part of the Halstead-Reitan neuropsychological battery (Reitan and Wolfson, 1993). More recently, modified versions of the TMT have become available, such as the Comprehensive Trail Making Test (Allen et al., 2012a) and the Trail Making subtest on the Delis-Kaplan Executive Function System (Fine et al., 2011; Allen et al., 2012b).

Here we utilize the Intermediate form of the TMT from the Halstead-Reitan neuropsychological test battery (Reitan and Wolfson, 1993). This form has two conditions. The first condition, *Trails A*, requires youth to connect fifteen encircled numbers in order, from 1 to 15, as quickly as possible. The second condition, *Trails B*, requires youth to alternate between connecting numbers and letters in order (i.e., 1-A-2-B and so on) as quickly as possible for a total of 15 connections. Performance on both Trails A and Trails B is thought to tap attention, psychomotor speed, and sequencing abilities. In addition, Trails B is thought to assess set-shifting, a commonly recognized executive function that requires individuals to switch their attention between two rules or tasks (Miyake et al., 2000). Often, investigators interested in studying the more "executive" components of the TMT focus on the difference in completion time for Trails B and Trails A. This difference is thought to partially account for the influence of baseline motoric speed or more basic cognitive abilities on performance and instead focus on the increased higher-order executive demands placed on participants during the Trails B condition, namely set-shifting. We will examine the neural correlates of this score (Trails B – A) as well as Trails B time directly.

The vast majority of studies examining the neural correlates of the TMT have been conducted with adults. To the best of our knowledge, only one study (Tamnes et al., 2010) has examined brain-behavior relations in typical youth using structural MRI and the TMT. Like the current study, these researchers utilized cortical thickness as their neuroanatomic phenotype; however, they directly correlated cortical thickness and TMT performance. As will be described in further detail below, our study examines how the coupling of cortical thickness values across the cortex vary as a function of TMT performance. Thus, the two studies use different analytic techniques to examine brain-behavior relations. In their study, Tamnes and colleagues examined cortical thickness and executive function correlations using the TMT and several other tasks. Surprisingly, the authors reported that most significant correlations between executive function task performance and cortical thickness were found in posterior brain regions. Only one task, another measure of set-shifting, called Plus Minus, was associated with precentral gyrus thickness. Thus, this study highlights the importance of non-frontal regions in accounting for individual differences in executive function in a pediatric sample.

Because of the scarcity of studies examining the neuroanatomic correlates of TMT or set-shifting in youth, we will turn to the adult literature to help generate hypotheses for our study. These investigations include studies of patients with lesions in different anatomic locations as well as both structural and functional magnetic resonance imaging (sMRI and fMRI, respectively) studies within the context of health, aging, and psychiatric illness. With regard to lesion studies, there is a large corpus of research implicating the frontal lobes in the completion of set-shifting tasks, including the TMT (Eslinger and Grattan, 1993; Stuss et al., 2001; Aron et al., 2004; McDonald et al., 2005; Yochim et al., 2007). However, the importance of the frontal lobes to task performance does not appear to be specific, as studies of patients with non-frontal lesions also demonstrate impairment on the TMT. In fact, a meta-analysis demonstrated that while frontal patients showed a small but statistically significant disadvantage on Trails A relative to patients with non-frontal lesions, a statistically significant disadvantage was *not* found for Trails B, as would be expected (Demakis, 2004). Thus, it is clear from this meta-analysis that damage to other brain regions results in impaired performance on this multifaceted task, consistent with both structural (Pa et al., 2010) and functional neuroimaging (Moll et al., 2002; Zakzanis et al., 2005; Jacobson et al., 2011) studies of typical and atypical populations, which are described in greater detail below.

Three fMRI studies conducted with healthy adults utilizing either a verbal adaptation of the TMT or a version with an MRI-safe stylus implicated the frontal lobes when comparing Trails B vs. A performance. Two of these studies (Moll et al., 2002; Zakzanis et al., 2005) specifically implicated the left dorsolateral prefrontal cortex, while the third study implicated the right inferior and middle frontal gyri (Jacobson et al., 2011) along with the right precentral gyrus. All of these studies also noted the involvement of posterior brain regions while completing the TMT (and in particular when the B vs. A conditions were compared). Moll et al. (2002) noted the involvement of the intraparietal sulcus bilaterally. Zakzanis et al. (2005) reported left middle and superior temporal gyri activation and right cingulate and paracentral lobule activity. Finally, Jacobson et al. (2011) reported involvement of the left middle temporal and angular gyri.

Thus, there appears to be support from structural imaging studies of typical and atypical adults, lesion studies, and functional imaging for the importance of both the frontal lobes and

posterior brain regions in the completion of the TMT. These findings fit with current thinking that different cognitive abilities are likely to be better understood from a functional (or structural) network perspective (for a review, see Park and Friston, 2013). Rather than focus on one modular region of the brain, network approaches suggest that it is the functioning of different clusters of brain regions that is important for higher-level cognition. Across studies, a number of different functional brain networks have been described, including the frontoparietal control, dorsal and ventral attention, somatosensory-motor, visual, language, and default mode networks (for a review, see Lee et al., 2012).

In an effort to add to this literature, the current study investigated how individual differences in structural covariance relate to TMT performance. Structural covariance refers to the observation that '…inter-individual differences in the structure of a brain region often covary with inter-individual differences in other brain regions (Alexander-Bloch et al., 2013a, p. 322). Our group has examined structural covariance using different methods, including graph analytic techniques (Alexander-Bloch et al., 2013b) and a method developed by Lerch et al. (2006) referred to as MACACC or Mapping Anatomical Correlations Across Cerebral Cortex. Using the latter technique, Lerch et al. demonstrated that cortical thickness correlation maps between a seed region in Broca's area and the rest of the cortex closely resembled white matter tractography maps generated from diffusion tensor imaging investigations. These findings suggested that correlations among regional gray matter measurements may indeed reflect the underlying white matter connectivity (and network structure of regions that are anatomically connected). Thus this technique is quite analogous to functional MRI, which relies on examining correlations among BOLD activation foci as a measure of functional connectivity. The MACACC technique has also identified structural covariance among regions implicated in highly replicated functional imaging networks, such as the default mode (Raznahan et al., 2011) and language (Lee et al., 2013) networks.

Furthermore, structural covariance has been found to be predictive of cognitive function (Lerch et al., 2006) and disease states (He et al., 2008). With regard to the former, Lerch and colleagues provided the first evidence that correlations among regional cortical thickness measurements index individual differences in intellectual abilities in typical youth. Following up on this, we investigated how individual differences in cortical thickness covariance related to vocabulary aptitude (Lee et al., 2013). Similar to Lerch's findings for intellectual abilities, we found that greater cortical thickness covariance among semantic hubs in the brain was related to higher scores on the Wechsler Vocabulary subtest (Lee et al., 2013).

In the current paper, we have chosen to examine cortical thickness covariance over the covariance of other measures of brain morphometry, such as regional surface area or gyrification, because prior work in our laboratory has demonstrated that individual differences in cortical thickness relate to variation in intellectual abilities (Shaw et al., 2006) as well as subclinical autistic and antisocial traits (Wallace et al., 2012). Thus, we applied a similar approach to the one used in Lee et al. (2013) to examine structural brain networks underpinning individual differences in TMT to test the hypothesis that stronger TMT performance will be associated with greater cross-cortical covariance in regions of cortex thought to be relevant to executive function abilities (e.g., the prefrontal cortex).

MATERIALS AND METHODS

PARTICIPANTS

The study's cross-sectional sample included 146 unrelated, typically-developing youth, ages 9–14 years, participating in an ongoing brain imaging study of single- and twin-birth children and adolescents being conducted in the Child Psychiatry Branch of the National Institute of Mental Health (NIMH; Giedd et al., 2009). The vast majority of participants were Caucasian ($n = 121$; 83%) and right-handed ($n = 128$; 88%). Data regarding age, IQ, and Trails Performance can be found in **Table 1**.

Inclusion criteria were as follows. Participants were required to: (a) be free of any developmental, learning, or psychiatric disorders as well any condition known to affect gross brain development; and (b) have provided useable data on both the TMT and a structural MRI scan (acquired on a GE 1.5 T scanner) that were acquired with 3 months of each other. [The vast majority (∼98%) of participants completed testing and scanning within the same week].

Verbal or written assent was obtained from minors along with written consent from the parents. The NIMH Institutional Review Board approved the protocol.

COGNITIVE MEASURES

Wechsler Intelligence Scales

The Wechsler Abbreviated Scale of Intelligence (WASI) was administered to all participants (Wechsler, 1999) as an estimate of overall intellectual abilities.

Trail Making Test

All participants completed the Intermediate form of the Halstead-Reitan TMT (Reitan and Wolfson, 1993). As stated earlier, participants are asked to draw lines between encircled numbers (Part A) or to alternate between connecting encircled numbers and letters arranged on a page (Part B) as quickly as they can. Because the focus of the current study was on relations between individual differences in performance and anatomical coupling, scores on the different TMT measures were age-standardized by regressing the effects of age out of raw scores (i.e., the time to complete Trails B or the difference in time between Trails B and Trails A) and saving the standardized residuals ($M = 0$; $SD = 1$). The two primary variables considered in the current study were the age-regressed standardized residuals of Trails B Time and the Difference between Trails B and Trails A Completion

Table 1 | Demographic information about the sample and mean TMT age-adjusted Z-scores.

	Age	IQ	Trails B (s)	Trails B Age Z-score	Trails B–A (s)	Trails B–A Age Z-score
M	12.23	114.28	28.50	−0.06	14.92	0.00
SD	1.80	12.08	10.66	0.91	9.08	1.00
Range	9–14	86–147	10–61	−1.76–2.82	0–48	−1.93–3.82

Time (Trails B–A). Note that lower Z-scores denote better (faster) performance.

Prior to conducting primary analyses, data were inspected for normality and outliers. Of the 153 eligible participants with both a useable scan and TMT data, seven were excluded due to being outliers (>3 *SD* from the mean) on Trails A, Trails B or the difference between Trails B and A. This resulted in the current sample of 146 participants.

MRI SCAN ACQUISITION AND PROCESSING METHODS

All MRI scans were acquired using the same General Electric 1.5 Tesla Signa Scanner at the National Institutes of Health Clinical Center in Bethesda, Maryland. Each participant contributed one scan. A three-dimensional spoiled gradient recalled echo sequence in the steady state, designed to optimize distinctions between gray matter, white matter, and cerebrospinal fluid was used to acquire 124 contiguous, 1.5-mm thick slices in the axial plane ($TE/TR = 5/24$ ms; flip angle = 45 degrees, matrix = 256 × 192, NEX = 1, FOV = 24 cm, acquisition time 9.9 min).

Montreal Neurological Institute's (MNI) automated CIVET pipeline was used for tissue classification and subsequent cortical thickness measurements. The native MRI scans were registered into standardized stereotaxic space and were corrected for non-uniformity artifacts (Sled et al., 1998) using a linear transformation (Collins et al., 1994). Tissue was classified into gray or white matter, spinal fluid, or background with a neural net classifier (Zijdenbos et al., 2002). Subsequently, the inner (white matter) and outer (pial) cortical surfaces were extracted using deformable surface-mesh models (MacDonald et al., 2000; Kim et al., 2005), and they were aligned non-linearly toward a standard template surface (Robbins et al., 2004).

Cortical thickness was quantified by measuring the linked distance between the white and pial surfaces (t-link metric) in native space (MacDonald et al., 2000; Lerch and Evans, 2005). A 30-mm surface-based diffusion-smoothing kernel (Chung et al., 2003) was utilized. These methods have been validated several ways. Validation methods include (a) manual measurements (Kabani et al., 2001), (b) population simulation (Lerch and Evans, 2005), and (c) validation within an Alzheimer's disease sample (Lerch et al., 2005).

All scans passed a two-stage quality assessment process which ensured the absence of (a) visible motion artifacts extending into the brain parenchyma in native images, and (b) visible errors in definition of the cortical ribbon based on an inspection of 3D reconstructions for the gray-white and pial surfaces in each scan. Furthermore, we graphically inspected the distribution of individual cortical thickness estimates within our sample at statistically significant peak foci to screen for outlier effects, as well as quantitatively tested for the lack of distorting outlier effects by re-running analyses after exclusion of any data point with a Cook's distance value of greater than 0.03. This value was calculated using the following formula: $d = 4/n - k - 1$ where n is the number of cases and k is the number of independent variables.

STATISTICAL ANALYSES

The method we employ for analysis of structural covariance requires regressing the effects of age and sex out of vertex-level cortical thickness measurements to prevent observed anatomical coupling being confounded by the effects of age and sex on separate brain regions (Lerch et al., 2006). Age terms that were removed from cortical thickness measurements included age and age-squared, consistent with the findings from our laboratory on the longitudinal trajectory of cortical gray matter development from childhood to young adulthood (Giedd et al., 1999).

In order to evaluate if children with better scores on the TMT demonstrate a greater degree of structural covariance (particularly in regions such as the prefrontal cortex), an estimate of the relatedness of cross-cortical vertex-based thickness was needed. Analysis of vertex-wise cortical thickness correlations with overall mean cortical thickness (MCT) provides a computationally efficient alternative to calculating and then summarizing all possible vertex-vertex correlations in the brain (Lerch et al., 2006). Therefore, in keeping with prior work (Raznahan et al., 2011; Lee et al., 2013), we examine vertex-MCT coupling as a proxy for the relatedness of each vertex with all other vertices. This approach permits examination of the interaction between MCT and TMT performance continuously using regression in the complete sample of 146 participants rather than requiring participants to be categorized into arbitrary categories of high vs. low performance.

For primary analyses, regression was used to predict vertex thickness at 40,962 points in each hemisphere using a package written for use in R statistics developed by colleagues at MNI. In particular, we sought to determine if the relationship between MCT and the thickness of a particular vertex varied as a function of TMT performance. Thus, we were most interested in identifying vertices in which there was an interaction between MCT and TMT performance. Regression equations to test for this interaction were as follows:

Trails B Time:

Cortical thickness$_{(\text{vertex j})}$ = Intercept + ß1(MCT[1]) + ß2 (Trails B time[2]) + ß3(MCT[1] * Trails B time[2]).

Difference in time for Trails B vs. Trails A:

Cortical thickness$_{(\text{vertex j})}$ = Intercept + ß1(MCT[1]) + ß2 (Trails B–A time[2]) + ß3(MCT[1] * Trails B–A time[2]).

A False Discovery Rate (FDR) adjustment (Benjamini and Hochberg, 1995) was applied to control for multiple comparisons (i.e., 40,962 regression analyses per hemisphere). Specifically, FDR-adjusted q-values were generated for all terms in the regression equation—that is, the main effect of MCT, main effect of Trails, and the MCT*TMT performance interaction. The FDR threshold applied was $q < 0.05$.

Exploratory age group analyses

Lastly, given that the focus of this special issue is on the *development* of executive functions in childhood, we ran exploratory analyses in order to begin to investigate if TMT-coupling relations vary as a function of age in childhood. We did this in two ways. First, we ran a linear regression predicting vertex-level cortical thickness using the following dependent variables: MCT

[1]The MCT measure utilized here was residualized, with the variance associated with age, age[2], and sex removed.

[2]The TMT measures utilized here (Trails B and the difference between Trails B and A) were residualized, with the variance associated age removed.

(age-standardized), TMT performance (age-standardized), age group (above or below the median age of 12.48) and their interactions (both two-way interactions and the three-way interaction). Regression equations used for these analyses are as follows.

Trails B Time:

Cortical thickness$_{(\text{vertex j})}$ = Intercept + ß1(MCT^1) + ß2 (Trails B time^2) + ß3 (Age Subgroup) + ß4 (MCT^1 * Trails B time^2) + ß5 (MCT^1 *Age Subgroup) + ß6 (Trails B time^2 *Age Subgroup) + ß7 (MCT^1 * Trails B time^2 *Age Subgroup).

Difference in time for Trails B vs. Trails A:

Cortical thickness$_{(\text{vertex j})}$ = Intercept + ß1(MCT^1) + ß2 (Trails B–A time^2) + ß3 (Age Subgroup) + ß4 (MCT^1 * Trails B–A time^2) + ß5 (MCT^1 *Age Subgroup) + ß6 (Trails B–A time^2 *Age Subgroup) + ß7 (MCT^1 * Trails B–A time^2 *Age Subgroup).

For these analyses, we were most interested in the three-way interaction for MCT*TMT*Age subgroup, as a significant interaction would suggest that the relations between anatomical coupling within the context of TMT performance varied as a function of age.

Second, we divided the sample into younger and older participants by splitting the group at the median age. We then re-ran the primary regression analyses in the younger and older samples to qualitatively compare the findings. This will be described in greater detail in the Results section.

RESULTS

In this manuscript, our primary research question was as follows: Is stronger TMT performance in childhood (as measured by time to complete Trails B and the difference between Trails B and A) associated with greater cross-cortical covariance in regions of cortex thought to be relevant to executive functions (e.g., the prefrontal cortex)? Stated another way, is the thickness of the prefrontal cortex and other cortical regions more highly correlated with the thickness of the rest of cortex (as estimated by MCT) in those with higher TMT scores?

This question was evaluated separately at every vertex in each hemisphere in the complete sample of 146 participants using the regression equations described above in the Materials and Methods section. In particular, we were interested in whether the MCT*TMT performance interaction was significant, as this would indicate that the strength of the relationship between MCT and a particular vertex's thickness varied as a function of TMT performance.

Regions in which statistically significant interactions were found between MCT and either Trails B Time (age-adjusted) or the Difference between Trails B and A (age-adjusted) are presented in **Figure 1**. Blue vertices are those in which the MCT*Trails B interaction was significant (following FDR correction, $q < 0.05$ for all terms in the regression equation), such that tighter correlations between MCT and the thickness of that vertex were found for those who were faster (better) on Trails B. Green vertices are those in which the MCT*Trails B vs. A Difference score interaction was significant, such that stronger coupling was found for those with better performance (i.e., smaller differences in time between Trails B and A). Vertices in red are those for which both of the regression equations' interaction terms were significant. Because the focus of the manuscript was on regions of the cortex in which MCT and vertex thickness correlations varied as a function of TMT performance, we have elected to leave the findings for main effects of MCT and TMT performance out of **Figure 1**. However, this information has been included in **Supplementary Figures S1** and **S2** for Trails B and Trails B–A, respectively.

As can be seen in **Figure 1**, a large swath of cortex in the superior and medial prefrontal cortex was more tightly coupled with the rest of the cortical ribbon in those who were faster at Trails B. When the difference between Trails B and A was considered, several smaller clusters of vertices were found to be more tightly coupled with the thickness of the rest of the cortex in better performers, including an overlapping region in the medial prefrontal cortex associated with better Trails B performance described above (in red in the figure). Additional regions included a cluster of vertices in dorsolateral prefrontal cortex, two small clusters near the temporal-parietal junction, and a cluster of vertices in superior parietal lobule (including a small region that overlapped with Trails B performance as shown in red). Lastly, there were also a few regions in which tighter coupling between MCT and vertex thickness was associated with poorer TMT performance. These results are summarized in **Supplementary Figure S3**.

To complement these analyses and demonstrate that the clusters of vertices displayed in **Figure 1** were associated with a greater degree of coupling with the rest of the cortex in those who were better performers (based on age-adjusted scores), we dichotomized the complete sample of 146 participants into those with scores in the lower and upper quartiles of the sample based on their age-adjusted Trails B score (or the difference in time between Trails B and A–age-adjusted). We then ran correlations between the thickness of the peak vertex identified in prior analyses and all vertices in the left and right hemisphere in the two groups—high/fast performers (those with scores in the lower quartile—denoting faster performance) and low/slow performers (those with scores in the upper quartile—denoting slower performance).

These findings are summarized for Trails B in **Figure 2** and for the difference between Trails B and A in **Figure 3**. For Trails B, the vertex in which the highest *t*-value was found for the interaction between MCT and TMT performance—referred to as the "peak vertex"—was identified in the medial prefrontal cortex (see **Figure 2A**). The thickness of this vertex was correlated with all other vertices in the high/fast and low/slow performing groups separately. The resulting correlation coefficients were projected onto the cortex and are presented in **Figure 2B**. In order to illustrate differences in the number of vertices that exceeded different correlation coefficient thresholds, the correlation range evaluated was truncated and **Figure 2C** presents the regions (and number of vertices) in which the correlation coefficients exceeded the following thresholds: $r > 0.1$, 0.3, 0.5, and 0.7. These values (i.e., number of vertices falling above and below the different thresholds) were compared for high/fast and low/slow performers utilizing chi-square. For all comparisons, the chi-square results were significant (all χ^2s > 100, ps < 0.001) in favor of the higher/faster performers having a greater proportion of vertices that exceeded the stated correlation coefficient threshold. Analogously, for the

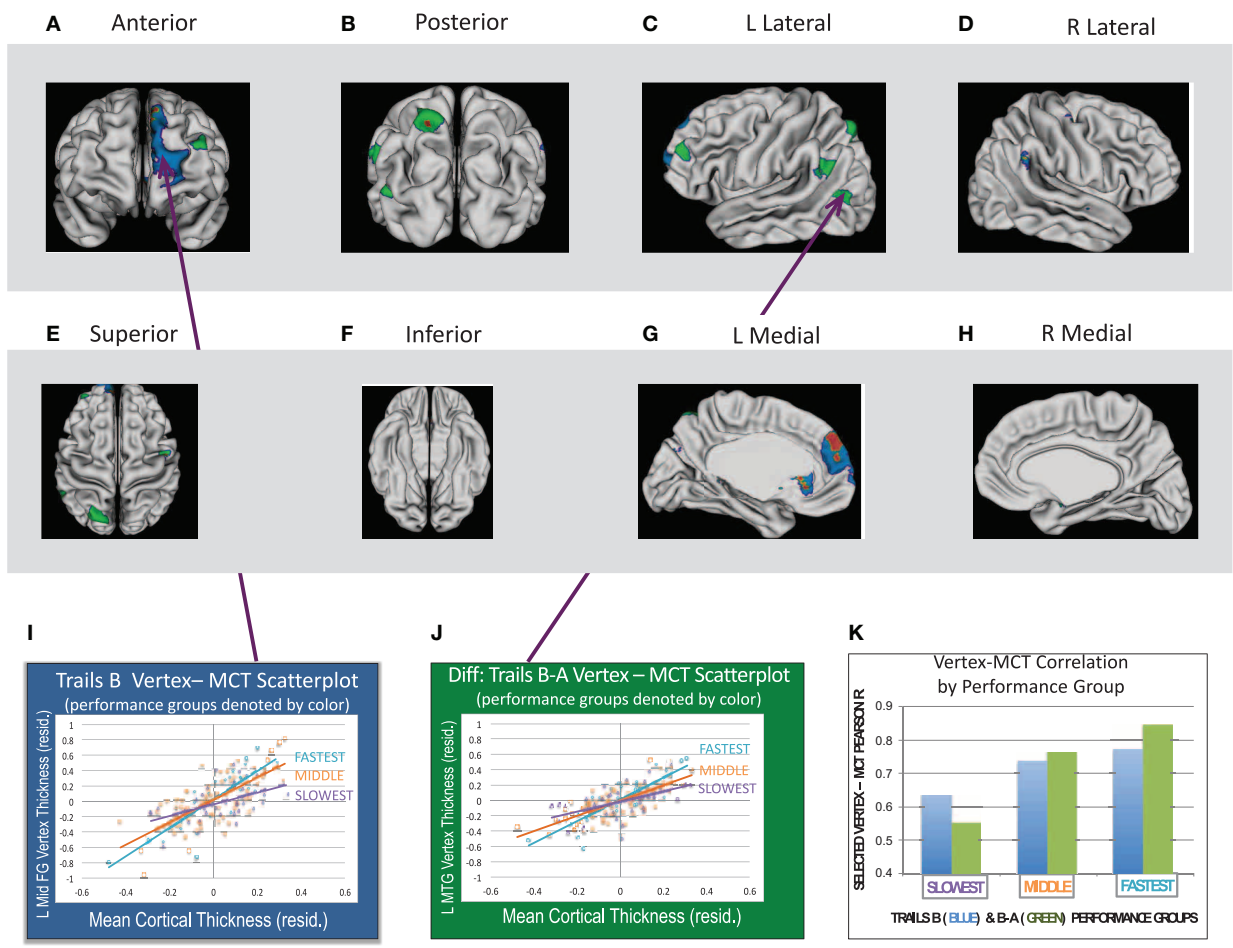


FIGURE 1 | Regions associated with greater cross-cortical coupling for those with stronger performance on Trails B, the Differences between Trails B and A, and Both Trails B and the B–A Difference. Two sets of linear regression analyses predicting cortical thickness at each vertex in both hemispheres were run in the complete sample ($n = 146$) of participants in order to evaluate if coupling between mean cortical thickness (MCT) and vertex thickness varied as a function of individual differences in either (1) Trails B time (age-adjusted) or (2) the difference in time between Trails B and A (age-adjusted). The regression equations were as follows. (1) For Trails B time: Cortical thickness $_{(\text{vertex j})}$ = Intercept + ß$_1$(MCT) + ß$_2$(Trails B time) + ß$_3$(MCT*Trails B). (2) For the Difference in time for Trails B vs. Trails A: Cortical thickness $_{(\text{vertex j})}$ = Intercept + ß$_1$(MCT^1) + ß$_2$(Trails B–A time) + ß$_3$(MCT*Trails B - A time). Note that for these analyses, the vertex-level dependent variables and MCT were age and sex standardized. (See Materials and Methods for details). The Trails B and B–A variables were age-standardized. *T*-statistics associated with the MCT*Trails interaction were corrected for multiple comparisons using a False Discovery Rate adjustment. Only those vertices with *T*s < -2.5 and *q*s < 0.05 are displayed in this figure in **(A–H)**. (Note that *t*-values are negative, because faster or shorter times are indicative of better performance). In these panels, blue vertices are those in which the MCT*Trails B interaction was significant, such that tighter coupling between MCT and the thickness of that vertex was found for those who were faster (better) on Trails B. Green vertices are those in which the MCT*Trails B vs. A Difference score interaction was significant, such that stronger coupling was found for those with better performance (i.e., smaller differences in time between Trails B and A). Vertices in red are those for which both of the regression equations' interaction terms were significant. **(I,J)** display relations between MCT and a vertex in the middle frontal gyrus (MNI coordinates = $x = -8$, $y = 68$, $z = 3$) or the middle temporal gyrus (MNI coordinates: $x = -47$, $y = -70$, $z = 16$), respectively, for performers stratified into three groups: the best/fastest performers shown in turquoise (those with scores in the lower quartile—denoting faster performance; $n = 37$), the middle performers in orange (middle 50% of sample; $n = 72$) and worst/slowest performers in purple (those with scores in the upper quartile—denoting slower performance; $n = 37$). As can be seen, a steeper regression line was associated with better performance. (Please note that performance was stratified into the three groups for illustrative purposes only here. The regression equations included a continuous measurement of performance on the TMT within the complete sample of 146 participants.) Lastly, **(K)** illustrates the Pearson *r* correlation coefficient values for (1) MCT and the selected vertex for Trails B and (2) MCT and the selected vertex for the difference between Trails B and A for the three subgroups included in the scatterplots shown in **(I,J)**. These values are shown for Trails B performance with the blue bars and the Difference between Trails B–A performance with the green bars. As can be seen, as performance group moves from slowest to fastest, the correlation between the pictured vertex and MCT increases.

difference between Trails B–A, the peak in the superior parietal lobule is used as the seed and the corresponding correlations are presented in **Figures 2B,C**. Lastly, for Trails B–A, correlation coefficient maps for the peak vertices in the middle and superior temporal lobe clusters and the dorsal and medial prefrontal cortex clusters are provided in **Supplementary Figure S4**.

EXPLORATORY ANALYSES EXAMINING THE EFFECTS OF AGE ON TMT-ANATOMICAL COUPLING FINDINGS

Given that the focus of the current special issue is on the development of executive functions, we undertook several exploratory analyses to evaluate differential age-effects on the TMT performance-coupling findings described above for the complete sample. Specifically, we ran a linear regression predicting vertex-level cortical thickness using the following dependent variables: MCT (age-standardized), TMT performance (age-standardized), age group (above or below the median age of 12.48) and their interactions (both two-way interactions and the three-way interaction). See the end of the Materials and Methods section for the equations utilized for Trails B and the difference between Trails B and A. For these analyses, we were most interested in the three-way MCT*TMT*Age group interaction, as this would suggest that the relations between anatomical coupling within the context of TMT performance varied as a function of age.

For Trails B, only three small regions were predicted significantly by the three-way interaction ($q < 0.05$). These included small regions in inferior medial prefrontal cortex, inferior somatosensory cortex, and the posterior cingulate. The approximate locations of these three regions are identified in **Supplementary Figure S5** with asterisks.

For these three-way interactions, results were such that tighter coupling was associated with better performance on Trails B (age-standardized) in the younger but not older subgroup. (For the older subgroup, the general trend in the data was for tighter coupling in the three regions being associated with *poorer* performance).

For the difference between Trails B and A, no statistically significant three-way interactions were identified, suggesting that the coupling-TMT performance findings were not modified by age. In addition to evaluating the occurrence of three-way interactions for Trails B and the difference between Trails B and A, we also divided our sample into two age-based subgroups (Younger: age less than the group median of 12.48; Older: age greater than or equal to the group median) in order to examine age effects in a more qualitative fashion. We then re-ran the initial regression equation used to answer the main study questions in these two subgroups: vertex thickness $\sim$ MCT + TMT + MCT*TMT performance. The MCT*TMT performance interaction results for the age-adjusted Trails B findings and the age-adjusted Trails B–A findings were projected onto the cortical surface in **Supplementary Figures S5**, **S6**, respectively. Cooler colors in these figures represent those in which the MCT*TMT performance interaction was significant in the whole sample, the younger subgroup, or both. In contrast, the warm colors represent regions in which the MCT*TMT interaction was significant for the older subgroup or both the older subgroup and the whole sample.

The results of the three-way MCT*TMT*Age interaction and the subgroup analyses suggest that age-effects on TMT-coupling are small within this limited age-range. However, these small effects suggest that younger age is associated with coupling among a greater number of cortical regions. This was tested by comparing the number of vertices that exceeded the FDR-corrected threshold ($q < 0.05$) for the MCT*TMT interaction for the younger and older subgroups. For both Trails B and the difference between Trials B and A, the chi-square findings were highly significant. For Trails B, 1229 vertices exceeded the threshold in the younger subgroup while only 297 exceeded this threshold in the older subgroup [$\chi^{2(1)} = 573$, $p < 0.001$]. Similarly, for the B–A difference, 800 vertices exceeded the threshold for the younger subgroup compared to 256 in the older subgroup [$\chi^{2(1)} = 281$, $p < 0.001$].

DISCUSSION

Adding to the literature on the neural correlates of executive function in childhood, here we demonstrate that individual differences on a commonly-administered executive function task, the Halstead-Reitan TMT, relate to the degree of anatomical coupling between the left prefrontal cortex and other distributed cortical regions. In particular, we found that for youth who were faster than their peers on Trails B (age-adjusted scores), there was greater coupling between a large swath of the prefrontal cortex, including portions of Brodmann areas 9 (dorsolateral prefrontal cortex) through 11 and the anterior cingulate, and the rest of cortex. When the difference between Trails B and A (age-adjusted) was considered, a network of mostly left-lateralized regions was found to be more strongly coupled with the rest of cortex, including clusters of vertices in the dorsolateral and dorsomedial prefrontal cortex, the posterior middle and superior temporal gyri (corresponding roughly to the angular and supramarginal gyri, respectively), and the superior parietal lobule.

These findings are the first to demonstrate how individual differences in structural (as opposed to functional) covariance relate to performance differences in executive functioning, a group of higher-level cognitive abilities that are believed to be important for academic outcomes (Blair and Razza, 2007) and are impaired in numerous developmental disorders (Ozonoff and Jensen, 1999). Despite the current study's focus on structural covariance, these findings are remarkably consistent with fMRI investigations into the functional correlates of TMT performance. Specifically, two studies (Moll et al., 2002; Zakzanis et al., 2005) implicated the left dorsolateral prefrontal cortex when Trails B performance was contrasted with Trails A. Furthermore, these two studies and a study conducted by Jacobson et al. (2011) reported the involvement of several posterior brain regions when Trails B activation was contrasted with Trails A activation. These included the intraparietal sulcus bilaterally (analogous to our supramarginal and angular gyri findings; Moll et al., 2002), the left middle and superior temporal gyri (Zakzanis et al., 2005; Jacobson et al., 2011), the angular gyrus (Jacobson et al., 2011), and the superior parietal lobule (for Trails B performance in particular; Allen et al., 2011).

Taken together, our structural covariance findings in concert with existing fMRI data provide support that a network of

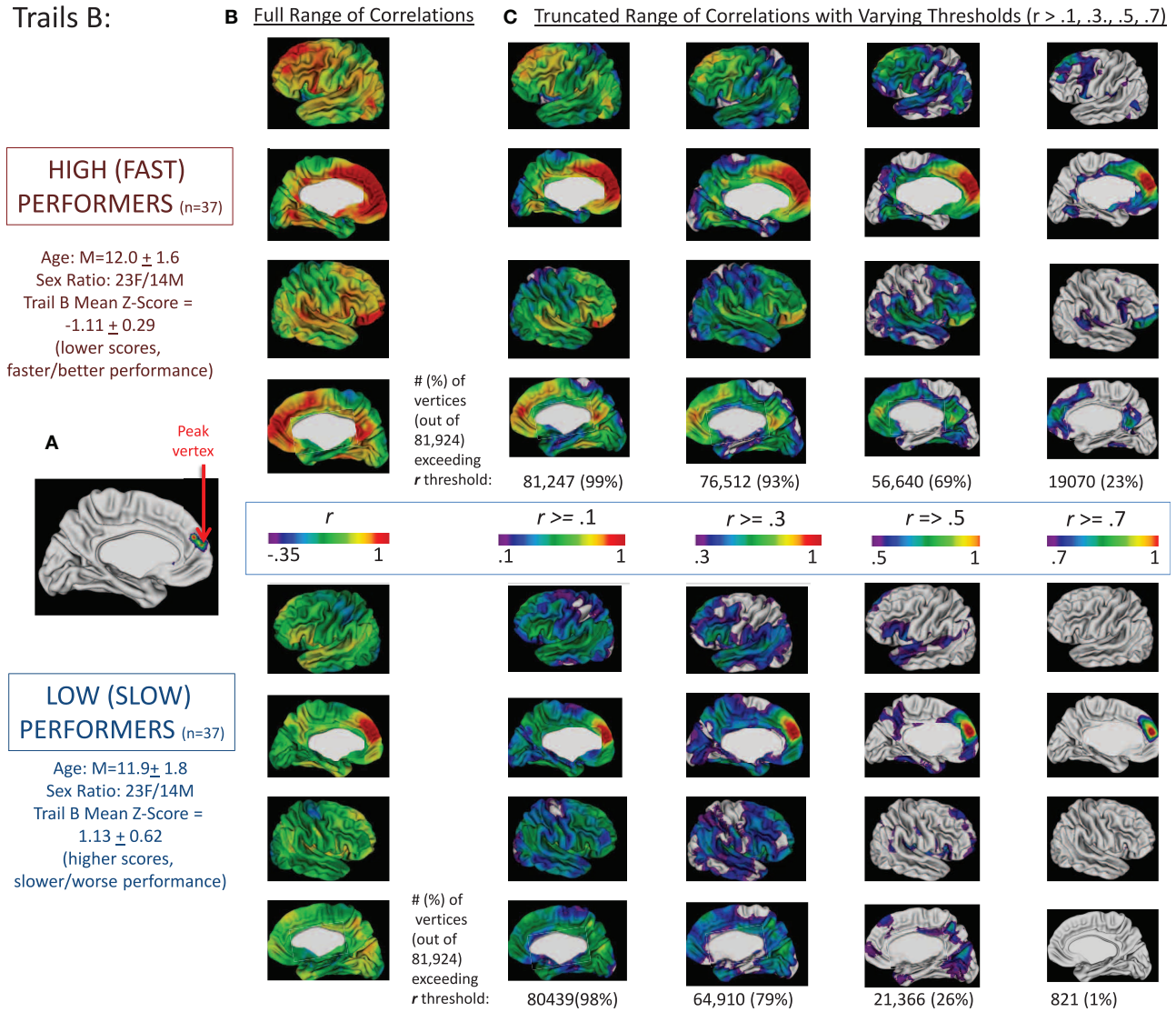


FIGURE 2 | Correlations between peak vertex thickness in the medial prefrontal cortex and the rest of the cortex in high ($n = 37$) and low ($n = 37$) performers on Trails B. The complete sample of 146 participants was divided into quartiles based on performance on Trails B (age-adjusted scores). We then ran two sets of correlations between the thickness of the peak vertex identified in prior analyses in medial prefrontal cortex [shown in **(A)**; MNI coordinates: $x = -9$, $y = 51$, $z = 14$] and all vertices in the left and right hemisphere for the group of high/fast performers (those with scores in the lower quartile—denoting faster performance; $n = 37$) and the low/slow performers (those with scores in the upper quartile—denoting slower performance; $n = 37$). The resulting correlation coefficients were projected onto the cortical sheet and are presented in **(B)**. In order to illustrate differences in the number of vertices that exceeded different correlation coefficient thresholds, the correlation range evaluated was truncated and **(C)** presents the regions (and number of vertices) in which the correlation coefficients exceeded the following thresholds: $r > 0.1$, 0.3, 0.5, and 0.7. These values (i.e., number of vertices falling above the different thresholds) are provided under the four images of the brains associated with each threshold for the high and low groups.

frontal and posterior brain regions is involved with successful TMT performance. Our study importantly extends the existing literature to include children for whom executive functioning abilities are developing. In addition, the current study's findings converge with a recent meta-analysis of adult lesion studies that strongly demonstrated that damage to brain regions other than the frontal lobes was just as likely to impair Trails B (and other executive test) performance as damage to the frontal lobes. Given the complexity of executive function tasks and the number of lower-level cognitive abilities that are involved (e.g., basic visual perception, focused attention, motor coordination and speed), it is not surprising that a network of regions working in unison is likely to underlie successful performance both in adulthood and childhood.

Analogous to our finding of greater structural covariance between the dorsolateral prefrontal cortex and the rest of the cortical ribbon, Cole and colleagues reported higher degrees of global functional connectivity in the lateral prefrontal cortex in

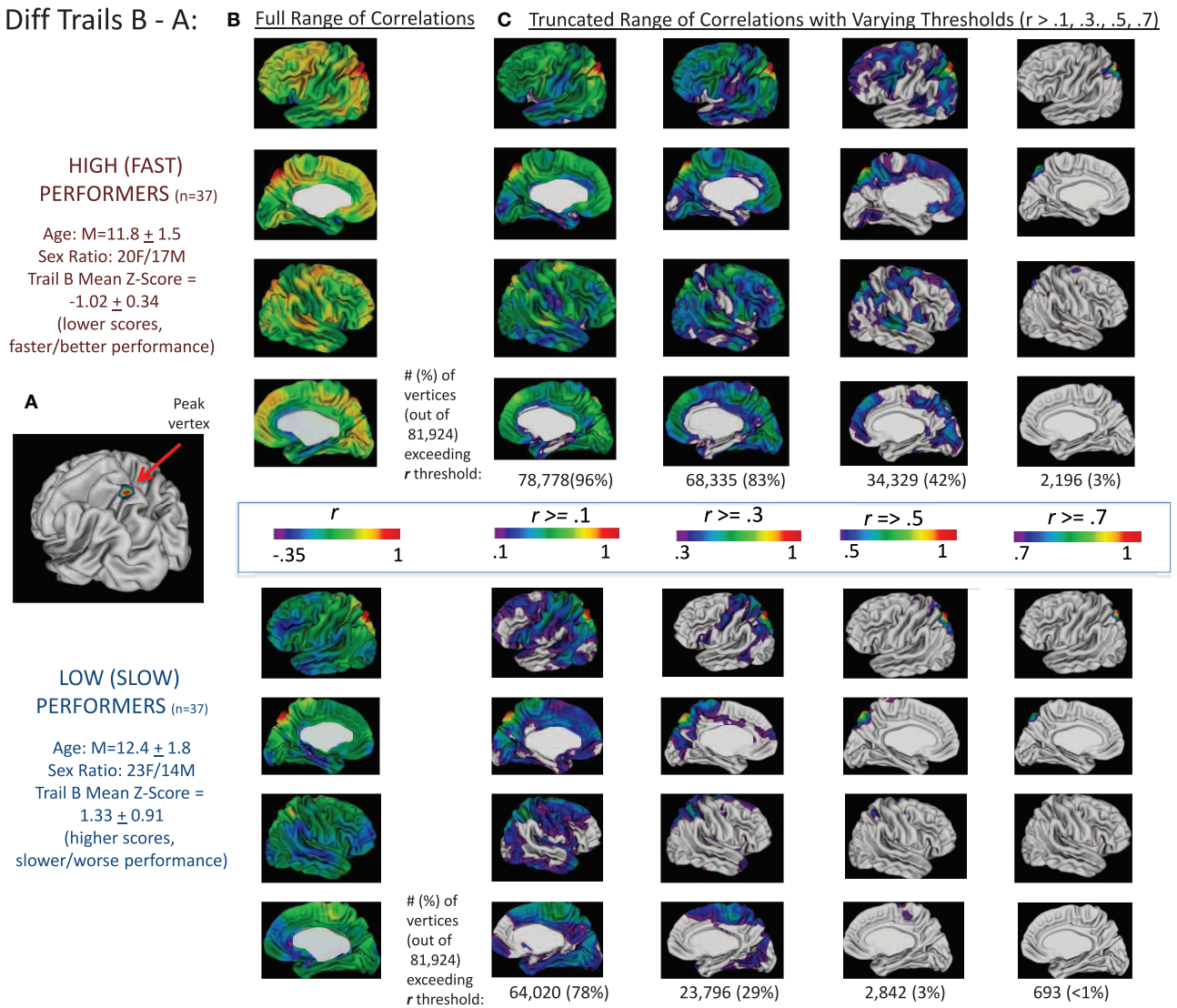


FIGURE 3 | Correlations between peak vertex thickness in the superior parietal lobule and the rest of the cortex in high ($n = 37$) and low ($n = 37$) performers on the Trails B–A difference score. Analogous to the procedures described in **Figure 2**, the complete sample of 146 participants was divided into quartiles based on performance on the difference in Trails B and A times (age-adjusted scores). We then ran two sets of correlations between the thickness of the peak vertex identified in prior analyses in superior parietal lobule [shown in **(A)**; MNI coordinates: $x = -22$, $y = -68$, $z = 62$] and all vertices in the left and right hemisphere for the group of high/fast performers (those with scores in the lower quartile—denoting faster performance; $n = 37$) and the low/slow performers (those with scores in the upper quartile—denoting slower performance; $n = 37$). The resulting correlation coefficients were projected onto the cortical sheet and are presented in **(B)**. In order to illustrate differences in the number of vertices that exceeded different correlation coefficient thresholds, the correlation range evaluated was truncated and **(C)** presents the regions (and number of vertices) in which the correlation coefficients exceeded the following thresholds: $r > 0.1$, 0.3, 0.5, and 0.7. These values (i.e., number of vertices falling above the different thresholds) are provided under the four images of the brains associated with each threshold for the high and low groups.

individuals with higher scores on measures of cognitive control (such as classic fluid intelligence tests; Cole et al., 2012). Moreover, an examination of the regions implicated in the current investigation of the TMT reveals an overlap with regions in the frontoparietal control and the default mode networks, two networks first described in functional connectivity studies (for a review, see Lee et al., 2012). In fact, it has been suggested that these networks are two of the most functionally connected in the brain (Cole et al., 2010). Thus, it is not surprising that higher degrees of structural covariance in these regions relates to higher performance on a complex, multifaceted executive function task, such as the TMT.

With regard to the examination of the impact of age on our TMT-coupling results, we found a small effect on these relations. However, the trend in our data tentatively suggested that anatomical coupling across *multiple* regions may be of greater importance for TMT success in younger participants, during a developmental period when executive function abilities are

rapidly developing. In contrast, as children and adolescents age, anatomical coupling in multiple regions may be less crucial for better performance. Instead, it could be that with age comes some regional specialization and greater reliance on cross-cortical coupling of a few select regions (e.g., reliance on the coupling of the dorsolateral prefrontal cortex in particular).

Given the importance of the prefrontal cortex in the current study and others examining executive functioning using different methodologies, we would be remiss if we did not focus some of our discussion on the importance of the frontal lobes to executive abilities in particular. In a review paper from 2001, Miller and Cohen provided an integrative theory about the functioning of the prefrontal cortex (Miller and Cohen, 2001). Based on a synthesis of neuroimaging, neurophysiological, anatomical, and computational investigations, they likened the prefrontal cortex to a "switch operator" in a rail system. Using this metaphor, they described the activity of the prefrontal cortex as a map that delineates which "tracks" or neural pathways are necessary for the completion of different cognitive tasks.

In this review, Miller and Cohen (2001) discussed the importance of the prefrontal cortex in maintaining "active representations" necessary to complete novel tasks requiring goal-directed behavior and flexibility. They suggested that one of the aspects of the prefrontal cortex that makes it unique is its ability to maintain active representations in the face of interference. Another unique feature of the prefrontal cortex is its high level of interconnectivity with sensory, motor, and limbic systems within the brain. These two qualities, among others, make the prefrontal cortex ideally-suited to serve as a "hub" and coordination center for higher-level cognitive abilities that require the work of multiple neuroanatomic regions.

In line with Miller and Cohen's conceptualization, more recent accounts of prefrontal cortex functioning such as the "gateway hypothesis" (Burgess et al., 2007) describe the rostral prefrontal cortex (roughly Brodmann area 10), an area implicated in the current investigation, as a "supervisory attentional gateway" that permits "stimulus-oriented" or "stimulus-independent" focused attention. These authors argue that the lateral rostral prefrontal cortex is more associated with the former, while the medial rostral prefrontal cortex is more associated with the latter. In the current investigation, both the lateral and medial prefrontal cortex were found to be more coupled in youth with higher TMT performance. Greater coupling in both of these regions certainly fits with TMT task demands—that is, one must attend to external stimuli (the encircled symbols on the page) and internal representations (maintaining a rote sequence of letters and numbers) in order to perform successfully on the task.

The current findings are also in line with the WHACH (what-how, abstract, cold-hot) model of prefrontal cortex functioning (O'Reilly, 2010). This model differentiates dorsal and ventral prefrontal functioning and suggests that the dorsal pathway is associated with guiding "how" to cope with information (i.e., "... transforming perception into action," p. 355) while the ventral pathway is associated with identifying "what" semantic information is relevant for a particular task (i.e., "... guiding the selection and retrieval of semantic/linguistic knowledge," p. 336). O'Reilly points out that the dorsal portions of the prefrontal cortex appear to be particularly relevant for transforming sensory inputs into motor outputs and for sequential ordering. These are two key aspects of successful TMT performance. Thus, higher coupling of the dorsal prefrontal cortex in better TMT performers provides additional support for the "how" conceptualization of dorsal prefrontal functioning.

Given the current study's findings and those of others, it may be that the prefrontal cortex represents a hub for higher-level executive abilities due to its inclusion in highly interconnected networks (dorsal portion of the frontoparietal control network and dorsal-medial portion of the default mode network). Based on the work of Buckner et al. (2009), it appears that all of the regions that were found to be more highly coupled in those with better TMT performance may indeed be locations of cortical hubs. Why might the "hub" regions implicated here be more coupled with the rest of cortex in youth who perform better on the TMT task? One possible explanation draws upon the Hebbian learning notion that neurons that "fire together, wire together" (Hebb, 1949). Thus, it may be that in youth who are better at executive tasks, the coordinated use of different regions of the brain, including the prefrontal cortex, results in a higher degree of anatomical coupling among these regions. An alternate explanation is that genetic factors contributing to the development of these brain regions are shared, and that youth who are better at these tasks are predisposed to more coordinated development in these regions.

The cross-sectional nature of this investigation precludes drawing any conclusions about these alternatives, a limitation of our study design. Another limitation of our study is that we focused on just one executive task, thus reducing the generalizability of our findings to other executive abilities. Furthermore, given the small age range of the sample studied here (9–14 years), we were only able to examine age-TMT-structural covariance relations in a preliminary way. A rigorous examination of age by performance effects on anatomical coupling, particularly within the context of a longitudinal study design, will be a crucial next step in understanding the complex unfolding of the development of executive abilities and how individual differences in performance emerge over time.

Despite these limitations, this study is the first of its kind to highlight the importance of structural covariance in relation to individual differences in executive function abilities in youth. Thus, it adds to the growing literature on the neural correlates of childhood executive functions and identifies neuroanatomic coupling as a biological substrate that may contribute to typical and atypical executive development. Consistent with fMRI connectivity work (Lee et al., 2012; Park and Friston, 2013), the present study demonstrates that successful performance on a multiply-determined executive function task is associated with greater anatomical coupling between the prefrontal cortex and other broadly-distributed cortical regions during childhood and adolescence. Thus, this study of individual differences in the context of typical development suggests that disorders of childhood associated with executive dysfunction (i.e., lower scores on tasks like the TMT), such as attention deficit hyperactivity disorder and autism spectrum disorder, might demonstrate more localized anatomical coupling in the frontal lobe and other regions.

AUTHOR CONTRIBUTIONS

Nancy Raitano Lee, Gregory L. Wallace, and Jay N. Giedd contributed to study design. Liv S. Clasen prepared data for analysis. Nancy Raitano Lee analyzed data and wrote the manuscript. Gregory L. Wallace, Liv S. Clasen, Armin Raznahan, and Jay N. Giedd critically revised the manuscript.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Mental Health (NCT00001246; Protocol ID 89-M-0006). We would like to thank the children and families who made this research possible.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00496/abstract

Supplementary Figure S1 | Regions in which there were main effects of mean cortical thickness, Trails B performance, and their interaction in predicting vertex thickness. This figure supplements **Figure 1** in the main document. Linear regression analyses predicting cortical thickness at each vertex in both hemispheres were run in the complete sample ($n = 146$) of participants in order to evaluate the effects of mean cortical thickness, Trails B age-adjusted scores, and their interaction. The regression equation was as follows: Cortical thickness $_{(\text{vertex j})}$ = Intercept + ß$_1$(MCT) + ß$_2$(Trails B time) + ß$_3$(MCT*Trails B). Note that for these analyses, the vertex-level dependent variables and MCT were age and sex standardized. (See Materials and Methods for details). The Trails B variables were age-standardized. *T*-statistics associated with each of the effects in the regression equation were corrected for multiple comparisons using a False Discovery Rate adjustment. Only those vertices with *q*s < 0.05 (associated with a *T*-threshold of 2.5) are displayed in this figure in **(A–H)**. Vertices in purple are those in which a main effect of MCT was found; vertices in blue are those in which main effects of MCT and Trails B were found such that thinner cortex was associated with better performance; vertices in turquoise green are those in which main effects of MCT and Trails B were found such that thicker cortex was associated with better performance; lastly, vertices in yellow are those in which an MCT*Trails B interaction was found such that greater coupling was associated with better performance.

Supplementary Figure S2 | Regions in which there were main effects of mean cortical thickness, Trails B–A performance, and their interaction in predicting vertex thickness. This figure also supplements **Figure 1** in the main document. Linear regression analyses predicting cortical thickness at each vertex in both hemispheres were run in the complete sample ($n = 146$) of participants in order to evaluate the effects of mean cortical thickness, Trails B–A age-adjusted scores, and their interaction. The regression equation was as follows: Cortical thickness $_{(\text{vertex j})}$ = Intercept + ß$_1$(MCT) + ß$_2$(Trails B–A time-age-adjusted) + ß$_3$(MCT*Trails B–A time- age-adjusted). Note that for these analyses, the vertex-level dependent variables and MCT were age and sex standardized. (See Materials and Methods for details). The Trails B–A variables were age-standardized. *T*-statistics associated with each of the effects in the regression equation were corrected for multiple comparisons using a False Discovery Rate adjustment. Only those vertices with *q*s < 0.05 (associated with a *T*-threshold of 2.5) are displayed in this figure in **(A–H)**. Vertices in purple are those in which a main effect of MCT was found; vertices in blue are those in which main effects of MCT and Trails B–A were found such that thinner cortex was associated with better performance; vertices in turquoise green are those in which main effects of MCT and Trails B–A were found such that thicker cortex was associated with better performance; lastly, vertices in yellow are those in which an MCT*Trails B–A interaction was found such that greater coupling was associated with better performance.

Supplementary Figure S3 | Regions associated with greater cross-cortical coupling for those with poorer performance on Trails B, the Differences between Trails B and A, and Both Trails B and the B–A Difference. This figure complements **Figure 1** in the main document in that it displays regions of the cortex in which greater coupling was associated with *poorer* performance. Two sets of linear regression analyses predicting cortical thickness at each vertex in both hemispheres were run in the complete sample ($n = 146$) of participants in order to evaluate if coupling between mean cortical thickness (MCT) and vertex thickness varied as a function of individual differences in either (1) Trails B time or (2) the difference in time between Trails B and A. The regression equations were as follows. (1) For Trails B time: Cortical thickness $_{(\text{vertex j})}$ = Intercept + ß$_1$(MCT) + ß$_2$(Trails B time) + ß$_3$(MCT*Trails B). (2) For the Difference in time for Trails B vs. Trails A: Cortical thickness $_{(\text{vertex j})}$ = Intercept + ß$_1$(MCT^1) + ß$_2$(Trails B–A time) + ß$_3$(MCT*Trails B − A time). Note that for these analyses, the vertex-level dependent variables and MCT were age and sex standardized. (See Materials and Methods for details). The Trails B and B–A variables were age-standardized. *T*-statistics associated with the MCT*Trails interaction were corrected for multiple comparisons using a False Discovery Rate adjustment. Only those vertices with *T*s $>$ 2.5 and *q*s < 0.05 are displayed in this figure in **(A–H)**. (Note that *t*-values are positive, because slower or longer times are indicative of poorer performance.) In these panels, blue vertices are those in which the MCT*Trails B interaction was significant, such that tighter coupling between MCT and the thickness of that vertex was found for those who were *slower* (worse) on Trails B. Green vertices are those in which the MCT*Trails B vs. A Difference score interaction was significant, such that stronger coupling was found for those with worse performance (i.e., greater differences in time between Trails B and A). Vertices in red are those for which both of the regression equations' interaction terms were significant. Lastly, **(I,J)** display relations between MCT and the thickness of two vertices in the inferior frontal gyrus—one associated with Trails B (MNI coordinates: $x = 25$, $y = 28$, $z = -17$) and the other associated with Trails B − A time (MNI coordinates: $x = 20$, $y = 26$, $z = -20$). Regression lines are for performers in three groups: the best/fastest performers shown in turquoise (those with scores in the lower quartile—denoting faster performance; $n = 37$), the middle performers in orange (middle 50% of sample; $n = 72$) and worst/slowest performers in purple (those with scores in the upper quartile—denoting slower performance; $n = 37$). As can be seen, a steeper regression line was associated with poorer performance. (Please note that performance was stratified into the three groups for illustrative purposes only here. The regression equations included a continuous measurement of performance on the TMT within the complete sample of 146 participants.)

Supplementary Figure S4 | Correlations between peak vertex thickness for clusters in the middle temporal, superior temporal, dorsolateral prefrontal, and dorsomedial prefrontal cortex in high ($n = 37$) and low ($n = 37$) performers on the Trails B–A difference score. Analogous to the procedures described in **Figure 3**, the complete sample of 146

participants was divided into quartiles based on performance on the difference in Trails B and A times (age-adjusted scores). We then ran two sets of correlations between the thickness of the peak vertex identified in prior analyses in the **(A)** posterior middle temporal gyrus, **(B)** posterior superior temporal gyrus, **(C)** dorsolateral prefrontal cortex, and **(D)** dorsomedial prefrontal cortex and all vertices in the left and right hemisphere for the group of high/fast performers (those with scores in the lower quartile—denoting faster performance; $n = 37$) and the low/slow performers (those with scores in the upper quartile—denoting slower performance; $n = 37$). The resulting correlation coefficients were projected onto the cortical sheet and can be viewed in **(A)** through **(D)**. Note: MNI coordinates for peaks included in this figure were as follows: **(A)** posterior middle temporal gyrus: $x = -47$, $y = -70$, $z = 16$; **(B)** posterior superior temporal gyrus: $x = -57$, $y = -54$, $z = 34$; **(C)** dorsolateral prefrontal cortex: $x = -39$, $y = 51$, $z = 16$; **(D)** dorsomedial prefrontal cortex: $x = -4$, $y = 53$, $z = 29$.

Supplementary Figure S5 | Regions in which an interaction between mean cortical thickness Trails B performance were found for younger participants ($n = 73$), older participants ($n = 73$), and the whole sample ($n = 146$). In order to investigate whether TMT-coupling relations vary as a function of age in childhood, we divided the sample into younger and older participants by splitting the group at the median age (12.48). We then re-ran the primary regression analyses [Cortical thickness $_{(\text{vertex j})}$ = Intercept + ß$_1$(MCT) + ß$_2$(Trails B time-age-adjusted) + ß$_3$(MCT*Trails B-age-adjusted)] in the younger ($n = 73$) and older ($n = 73$) subgroups to qualitatively compare the findings. Note that for these analyses, the vertex-level dependent variables and MCT were age and sex standardized. (See Materials and Methods for details). The Trails B variables were age-standardized. Vertices exceeding the FDR-adjusted threshold for MCT*Trails B interaction ($q < 0.05$; T-threshold of 2.5) were projected onto the cortex using the following color code. (1) Vertices associated with a statistically significant MCT*Trails B interaction in the whole sample were coded royal blue; (2) vertices that were only found to be statistically significant in the younger sample were coded turquoise blue; (3) vertices associated with a statistically significant interaction in analyses of both the whole sample and the younger sample were coded green; (4) vertices with statistically significant interactions in the older subgroup were coded orange; and (5) vertices with statistically significant interaction terms in both analyses of the older subgroup and the whole sample were coded red. In order to evaluate age-effects using a more rigorous technique statistically, we also ran a series of linear regression analyses predicting vertex-level cortical thickness using the following equation: Cortical thickness$_{(\text{vertex j})}$ = Intercept + ß1(MCT) + ß2(Trails B time) + ß3 (Age Subgroup) + ß4 (MCT*Trails B time-age adjusted) + ß5 (MCT*Age Subgroup) + ß6 (Trails B time-age adjusted *Age Subgroup) + ß7 (MCT*Trails B time-age adjusted*Age Subgroup). Again note that for these analyses, the vertex-level dependent variables and MCT were age and sex standardized (see Materials and Methods for details) and the Trails B variables were age-standardized. Regions with an asterisk (*) denote the approximate location of the statistically significant three-way interactions that survived the FDR adjusted T-value of 2.7 ($q < 0.05$). For the three regions in which there was a three-way interaction, the findings were such that greater coupling was associated with better Trails B (age-adjusted) performance in younger participants. The general trend in the data was for the opposite to be true in the older subgroup—that is tighter coupling was associated with poorer performance.

Supplementary Figure S6 | Regions in which an interaction between mean cortical thickness Trails B–A performance were found for younger participants ($n = 73$), older participants ($n = 73$), and the whole sample ($n = 146$). In order to investigate whether TMT-coupling relations vary as a function of age in childhood, we divided the sample into younger and older participants by splitting the group at the median age (12.44). We then re-ran the primary regression analyses [Cortical thickness $_{(\text{vertex j})}$ = Intercept + ß$_1$(MCT) + ß$_2$(Trails B–A time-age-adjusted) + ß$_3$(MCT*Trails B–A-age-adjusted)] in the younger ($n = 73$) and older ($n = 73$) subgroups to qualitatively compare the findings. Vertices with interaction terms exceeding the FDR-adjusted threshold for regression analyses ($T > 2.5$; $q < 0.05$) were projected onto the cortex using the following color code. (1) Vertices associated with statistically significant MCT*Trails B–A interaction terms in the whole sample were coded royal blue; (2) vertices that were only found to be statistically significant in the younger sample were coded turquoise blue; (3) vertices associated with a statistically significant interaction in analyses of both the whole sample and the younger sample were coded green; (4) vertices with statistically significant interactions in the older subgroup were coded orange; and (5) vertices with statistically significant interaction terms in both analyses of the older subgroup and the whole sample were coded red. In order to evaluate age-effects using a more rigorous technique statistically, we also ran a series of linear regression analyses predicting vertex-level cortical thickness using the following equation: Cortical thickness$_{(\text{vertex j})}$ = Intercept + ß1(MCT) + ß2(Trails B–A time) + ß3 (Age Subgroup) + ß4 (MCT*Trails B–A time-age-adjusted) + ß5 (MCT*Age Subgroup) + ß6 (Trails B–A time-age adjusted *Age Subgroup) + ß7 (MCT*Trails B time-age adjusted*Age Subgroup). Again note that for these analyses, the vertex-level dependent variables and MCT were age and sex standardized (see Materials and Methods for details) and the Trails B variables were age-standardized. Unlike Trails B performance, no statistically significant three-way interactions were found (all qs > 0.05).

REFERENCES

Alexander-Bloch, A., Giedd, J. N., and Bullmore, E. (2013a). Imaging structural co-variance between human brain regions. *Nat. Rev. Neurosci.* 14, 322–336. doi: 10.1038/nrn3465

Alexander-Bloch, A., Raznahan, A., Bullmore, E., and Giedd, J. (2013b). The convergence of maturational change and structural covariance in human cortical networks. *J. Neurosci.* 33, 2889–2899. doi: 10.1523/JNEUROSCI.3554-12.2013

Allen, D. N., Thaler, N. S., Barchard, K. A., Vertinski, M., and Mayfield, J. (2012a). Factor structure of the Comprehensive Trail Making Test in children and adolescents with brain dysfunction. *Psychol. Assess.* 24, 964–972. doi: 10.1037/a0028521

Allen, D. N., Thaler, N. S., Ringdahl, E. N., Barney, S. J., and Mayfield, J. (2012b). Comprehensive Trail Making Test performance in children and adolescents with traumatic brain injury. *Psychol. Assess.* 24, 556–564. doi: 10.1037/a0026263

Allen, M. D., Owens, T. E., Fong, A. K., and Richards, D. R. (2011). A functional neuroimaging analysis of the Trail Making Test-B: implications for clinical application. *Behav. Neurol.* 24, 159–171. doi: 10.1155/2011/476893

Aron, A. R., Monsell, S., Sahakian, B. J., and Robbins, T. W. (2004). A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. *Brain* 127(Pt 7), 1561–1573. doi: 10.1093/brain/awh169

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300. doi: 10.2307/2346101

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Braver, T. S., and Ruge, H. (2006). "Neuroimaging of executive functions," in *Handbook of Functional Neuroimaging of Cognition*, *2nd Edn.*, eds R. Cabeza and A. Kingstone (Cambridge, MA: MIT Press), 307–348.

Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., et al. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J. Neurosci.* 29, 1860–1873. doi: 10.1523/JNEUROSCI.5062-08.2009

Burgess, P. W., Dumontheil, I., and Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends Cogn. Sci.* 11, 290–298. doi: 10.1016/j.tics.2007.05.004

Chung, M. K., Worsley, K. J., Robbins, S., Paus, T., Taylor, J., Giedd, J. N., et al. (2003). Deformation-based surface morphometry applied to gray matter deformation. *Neuroimage* 18, 198–213. doi: 10.1016/S1053-8119(02)00017-4

Cole, M. W., Pathak, S., and Schneider, W. (2010). Identifying the brain's most globally connected regions. *Neuroimage* 49, 3132–3148. doi: 10.1016/j.neuroimage.2009.11.001

Cole, M. W., Yarkoni, T., Repovs, G., Anticevic, A., and Braver, T. S. (2012). Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J. Neurosci.* 32, 8988–8999. doi: 10.1523/JNEUROSCI.0536-12.2012

Collins, D. L., Neelin, P., Peters, T. M., and Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.* 18, 192–205. doi: 10.1097/00004728-199403000-00005

Conklin, H. M., Luciana, M., Hooper, C. J., and Yarger, R. S. (2007). Working memory performance in typically developing children and adolescents: behavioral evidence of protracted frontal lobe development. *Dev. Neuropsychol.* 31, 103–128. doi: 10.1207/s15326942dn3101_6

Demakis, G. J. (2004). Frontal lobe damage and tests of executive processing: a meta-analysis of the category test, stroop test, and trail-making test. *J. Clin. Exp. Neuropsychol.* 26, 441–450. doi: 10.1080/13803390490510149

Eslinger, P. J., and Grattan, L. M. (1993). Frontal lobe and frontal-striatal substrates for different forms of human cognitive flexibility. *Neuropsychologia* 31, 17–28. doi: 10.1016/0028-3932(93)90077-D

Fine, E. M., Delis, D. C., and Holdnack, J. (2011). Normative adjustments to the D-KEFS trail making test: corrections for education and vocabulary level. *Clin. Neuropsychol.* 25, 1331–1344. doi: 10.1080/13854046.2011.609838

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nat. Neurosci.* 2, 861–863. doi: 10.1038/13158

Giedd, J. N., Lalonde, F. M., Celano, M. J., White, S. L., Wallace, G. L., Lee, N. R., et al. (2009). Anatomical brain magnetic resonance imaging of typically developing children and adolescents. *J. Am. Acad. Child Adolesc. Psychiatry* 48, 465–470. doi: 10.1097/CHI.0b013e31819f2715

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179. doi: 10.1073/pnas.0402680101

He, Y., Chen, Z., and Evans, A. (2008). Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. *J. Neurosci.* 28, 4756–4766. doi: 10.1523/JNEUROSCI.0141-08.2008

Hebb, D. O. (1949). *The Organization of Behavior*. New York, NY: Wiley and Sons.

Hooper, C. J., Luciana, M., Conklin, H. M., and Yarger, R. S. (2004). Adolescents' performance on the Iowa Gambling Task: implications for the development of decision making and ventromedial prefrontal cortex. *Dev. Psychol.* 40, 1148–1158. doi: 10.1037/0012-1649.40.6.1148

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Jacobson, S. C., Blanchard, M., Connolly, C. C., Cannon, M., and Garavan, H. (2011). An fMRI investigation of a novel analogue to the Trail-Making Test. *Brain Cogn.* 77, 60–70. doi: 10.1016/j.bandc.2011.06.001

Kabani, N., Le Goualher, G., MacDonald, D., and Evans, A. C. (2001). Measurement of cortical thickness using an automated 3-D algorithm: a validation study. *Neuroimage* 13, 375–380. doi: 10.1006/nimg.2000.0652

Kim, J. S., Singh, V., Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., et al. (2005). Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *Neuroimage* 27, 210–221. doi: 10.1016/j.neuroimage.2005.03.036

Lee, M. H., Hacker, C. D., Snyder, A. Z., Corbetta, M., Zhang, D., Leuthardt, E. C., et al. (2012). Clustering of resting state networks. *PLoS ONE* 7:e40370. doi: 10.1371/journal.pone.0040370

Lee, N. R., Raznahan, A., Wallace, G. L., Alexander-Bloch, A., Clasen, L. S., Lerch, J. P., et al. (2013). Anatomical coupling among distributed cortical regions in youth varies as a function of individual differences in vocabulary abilities. *Hum. Brain Mapp.* 35, 1885–1895. doi: 10.1002/hbm.22299

Lerch, J. P., and Evans, A. C. (2005). Cortical thickness analysis examined through power analysis and a population simulation. *Neuroimage* 24, 163–173. doi: 10.1016/j.neuroimage.2004.07.045

Lerch, J. P., Pruessner, J. C., Zijdenbos, A., Hampel, H., Teipel, S. J., and Evans, A. C. (2005). Focal decline of cortical thickness in Alzheimer's disease identified by computational neuroanatomy. *Cereb. Cortex* 15, 995–1001. doi: 10.1093/cercor/bhh200

Lerch, J. P., Worsley, K., Shaw, W. P., Greenstein, D. K., Lenroot, R. K., Giedd, J., et al. (2006). Mapping anatomical correlations across cerebral cortex (MACACC) using cortical thickness from MRI. *Neuroimage* 31, 993–1003. doi: 10.1016/j.neuroimage.2006.01.042

Lezak, M. D., Howieson, D. N., and Loring, D. W. (2004). *Neuropsychological Assessment, 4th Edn*. New York, NY: Oxford University Press.

Luciana, M., Conklin, H. M., Hooper, C. J., and Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Dev.* 76, 697–712. doi: 10.1111/j.1467-8624.2005.00872.x

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., and Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Dev.* 75, 1357–1372. doi: 10.1111/j.1467-8624.2004.00745.x

MacDonald, D., Kabani, N., Avis, D., and Evans, A. C. (2000). Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage* 12, 340–356. doi: 10.1006/nimg.1999.0534

McDonald, C. R., Delis, D. C., Norman, M. A., Tecoma, E. S., and Iragui-Madozi, V. I. (2005). Is impairment in set-shifting specific to frontal-lobe dysfunction? Evidence from patients with frontal-lobe or temporal-lobe epilepsy. *J. Int. Neuropsychol. Soc.* 11, 477–481. doi: 10.1017/S1355617705050484

Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moll, J., de Oliveira-Souza, R., Moll, F. T., Bramati, I. E., and Andreiuolo, P. A. (2002). The cerebral correlates of set-shifting: an fMRI study of the trail making test. *Arq. Neuropsiquiatr.* 60, 900–905. doi: 10.1590/S0004-282X2002000600002

O'Reilly, R. C. (2010). The What and How of prefrontal cortical organization. *Trends Neurosci.* 33, 355–361. doi: 10.1016/j.tins.2010.05.002

Orellana, G., and Slachevsky, A. (2013). Executive functioning in schizophrenia. *Front. Psychiatry* 4:35. doi: 10.3389/fpsyt.2013.00035

Ozonoff, S., and Jensen, J. (1999). Brief report: specific executive function profiles in three neurodevelopmental disorders. *J. Autism Dev. Disord.* 29, 171–177. doi: 10.1023/A:1023052913110

Pa, J., Possin, K. L., Wilson, S. M., Quitania, L. C., Kramer, J. H., Boxer, A. L., et al. (2010). Gray matter correlates of set-shifting among neurodegenerative disease, mild cognitive impairment, and healthy older adults. *J. Int. Neuropsychol. Soc.* 16, 640–650. doi: 10.1017/S1355617710000408

Park, H. J., and Friston, K. (2013). Structural and functional brain networks: from connections to cognition. *Science* 342, 1238411. doi: 10.1126/science.1238411

Partington, J. E., and Leiter, R. G. (1949). Partington's pathway test. *Psychol. Service Center J.* 1, 11–20.

Pennington, B. F., and Ozonoff, S. (1996). Executive functions and developmental psychopathology. *J. Child Psychol. Psychiatry* 37, 51–87. doi: 10.1111/j.1469-7610.1996.tb01380.x

Raznahan, A., Lerch, J. P., Lee, N., Greenstein, D., Wallace, G. L., Stockman, M., et al. (2011). Patterns of coordinated anatomical change in human cortical development: a longitudinal neuroimaging study of maturational coupling. *Neuron* 72, 873–884. doi: 10.1016/j.neuron.2011.09.028

Reitan, R. M., and Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation, 2nd Edn*. Tuscon, AZ: Neuropsychology Press.

Robbins, S., Evans, A. C., Collins, D. L., and Whitesides, S. (2004). Tuning and comparing spatial normalization methods. *Med. Image Anal.* 8, 311–323. doi: 10.1016/j.media.2004.06.009

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., et al. (2006). Intellectual ability and cortical development in children and adolescents. *Nature* 440, 676–679. doi: 10.1038/nature04513

Sled, J. G., Zijdenbos, A. P., and Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* 17, 87–97. doi: 10.1109/42.668698

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychol. Bull.* 139, 81–132. doi: 10.1037/a0028727

Stuss, D. T., Bisschop, S. M., Alexander, M. P., Levine, B., Katz, D., and Izukawa, D. (2001). The Trail Making Test: a study in focal lesion patients. *Psychol. Assess.* 13, 230–239. doi: 10.1037/1040-3590.13.2.230

Tamnes, C. K., Ostby, Y., Fjell, A. M., Westlye, L. T., Due-Tonnessen, P., and Walhovd, K. B. (2010). Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cereb. Cortex* 20, 534–548. doi: 10.1093/cercor/bhp118

Wallace, G. L., Shaw, P., Lee, N. R., Clasen, L. S., Raznahan, A., Lenroot, R. K., et al. (2012). Distinct cortical correlates of autistic versus antisocial traits in a longitudinal sample of typically developing youth. *J. Neurosci.* 32, 4856–4860. doi: 10.1523/JNEUROSCI.6214-11.2012

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.

Welsh, M. C., and Pennington, B. F. (1988). Assessing frontal-lobe functioning in children - views from developmental-psychology. *Dev. Neuropsychol.* 4, 199–230. doi: 10.1080/87565648809540405

Yochim, B., Baldo, J., Nelson, A., and Delis, D. C. (2007). D-KEFS Trail Making Test performance in patients with lateral prefrontal cortex lesions. *J. Int. Neuropsychol. Soc.* 13, 704–709. doi: 10.1017/S1355617707070907

Zakzanis, K. K., Mraz, R., and Graham, S. J. (2005). An fMRI study of the Trail Making Test. *Neuropsychologia* 43, 1878–1886. doi: 10.1016/j.neuropsychologia.2005.03.013

Zelazo, P. D., and Muller, U. (2002). "Executive function in typical and atypical development," in *Handbook of Childhood Cognitive Development,* ed U. Goswami (Oxford: Blackwell), 445–469. doi: 10.1002/9780470996652.ch20

Zijdenbos, A. P., Forghani, R., and Evans, A. C. (2002). Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE Trans. Med. Imaging* 21, 1280–1291. doi: 10.1109/TMI.2002.806283

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 February 2014; accepted: 06 May 2014; published online: 01 July 2014.
Citation: Lee NR, Wallace GL, Raznahan A, Clasen LS and Giedd JN (2014) Trail making test performance in youth varies as a function of anatomical coupling between the prefrontal cortex and distributed cortical regions. Front. Psychol. ***5****:496. doi: 10.3389/fpsyg.2014.00496*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Lee, Wallace, Raznahan, Clasen and Giedd. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 24 June 2014
doi: 10.3389/fpsyg.2014.00655

The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective

Gina Austin*, Karoline Groppe and Birgit Elsner

Developmental Psychology, Department of Psychology, University of Potsdam, Potsdam, Germany

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Gary Morgan, City University London, UK
Daniela Kloo, University of Salzburg, Austria

****Correspondence:***
Gina Austin, Developmental Psychology, Department of Psychology, University of Potsdam, Karl-Liebknecht-Street 24/25, 14476 Potsdam, Germany
e-mail: gina.austin@uni-potsdam.de

There is robust evidence showing a link between executive function (EF) and theory of mind (ToM) in 3- to 5-year-olds. However, it is unclear whether this relationship extends to middle childhood. In addition, there has been much discussion about the nature of this relationship. Whereas some authors claim that ToM is needed for EF, others argue that ToM requires EF. To date, however, studies examining the longitudinal relationship between distinct subcomponents of EF [i.e., attention shifting, working memory (WM) updating, inhibition] and ToM in middle childhood are rare. The present study examined (1) the relationship between three EF subcomponents (attention shifting, WM updating, inhibition) and ToM in middle childhood, and (2) the longitudinal reciprocal relationships between the EF subcomponents and ToM across a 1-year period. EF and ToM measures were assessed experimentally in a sample of 1,657 children (aged 6–11 years) at time point one (t1) and 1 year later at time point two (t2). Results showed that the concurrent relationships between all three EF subcomponents and ToM pertained in middle childhood at t1 and t2, respectively, even when age, gender, and fluid intelligence were partialled out. Moreover, cross-lagged structural equation modeling (again, controlling for age, gender, and fluid intelligence, as well as for the earlier levels of the target variables), revealed partial support for the view that early ToM predicts later EF, but stronger evidence for the assumption that early EF predicts later ToM. The latter was found for attention shifting and WM updating, but not for inhibition. This reveals the importance of studying the exact interplay of ToM and EF across childhood development, especially with regard to different EF subcomponents. Most likely, understanding others' mental states at different levels of perspective-taking requires specific EF subcomponents, suggesting developmental change in the relations between EF and ToM across childhood.

Keywords: executive function, theory of mind, longitudinal, middle childhood, attention shifting, inhibition, working memory updating

INTRODUCTION

A major achievement of early development occurs when a child is able to impute mental states to himself/herself and others in order to predict and explain behavior ("theory of mind," ToM; Frith and Frith, 1999). This ability enables the individual to function in social groups and thus constitutes a crucial aspect of social competence. The development of ToM starts when the child is an infant and continues right the way through to the adolescent years (Lalonde and Chandler, 2002).

A critical test for ToM is the *first-order false-belief task* which is mastered at around the age of 4 years. One classical task requires the child to infer which belief a character in a story has about the location of an object which has been hidden in the character's presence, and has then been hidden somewhere else without the character knowing this (Wimmer and Perner, 1983). As children progress through childhood, they are able to solve more complex, so-called higher-order ToM tasks. One of the most commonly used is the *second-order false-belief task* (Perner and Wimmer, 1985) which requires the child to infer a story character's belief about another person's belief. The age of mastering second-order false-belief ranges from about 6–7 years, depending on the sample and method used (for a review see Miller, 2009).

Several related abilities of ToM have been identified of which executive function (EF) in particular has received considerable investigation and has led to much theoretical discussion. EF refers to an array of different processes relating to self-control. They develop in the preschool years and continue to do so right up to adolescence (Zelazo and Carlson, 2012). These processes enable the control of thought, action, and emotion, and they include overlapping but distinct EF subcomponents such as attention shifting, inhibition, and updating of working memory (WM updating; Miyake et al., 2000). As regards these specific EF subcomponents, ToM (first-order false-belief) understanding appears to require the ability to suppress one's own knowledge about reality (inhibition) in order to be able to put oneself into the shoes of the other (attention shifting) and then actively hold the key elements of the story in mind where this information can be monitored and updated in order to make an inference (WM updating; Doherty, 2009).

There is robust evidence that links ToM, (especially first-order false-belief) and these aforementioned specific EF subcomponents – including inhibition (Hughes, 1998a; Carlson and Moses, 2001; Flynn et al., 2004), attention shifting (Frye et al., 1995; Hughes, 1998a), and WM updating (Davis and Pratt, 1995; Keenan et al., 1998) in children aged 3–5 years (see Perner and Lang, 1999, for a review). Several reasons for this relationship have been put forward. For example, EF and ToM make major developments during the preschool years, they seem to share a common neurological basis (prefrontal cortex), and individuals suffering from autism show deficiencies in both (Carlson and Moses, 2001; Hill, 2004).

In need of clarification is whether the robust relationship between EF and ToM found in preschoolers extends to older children. Although there is some evidence that more advanced EF and ToM measures do show positive associations, more studies are needed to confirm this (Miller, 2009). For instance, it has been found that the EF–ToM relationship extends to children between $4^1/_2$ and $6^1/_2$ years for second-order false-belief tasks and more demanding EF tasks (Perner et al., 2002). Similar results have been reported for children of middle childhood (Yang et al., 2009; Calderon et al., 2010) and in adolescents (Vetter et al., 2013). However, results in a sample of $8^1/_2$ year-olds with and without attention deficit hyperactivity disorder (ADHD) were less conclusive (Charman et al., 2001). A correlation between EF and ToM was found in the control group (typically developing children) but as soon as age and intelligence were partialled out, the two constructs were no longer significantly correlated. These inconsistent results show that more studies are needed to clarify the relationship between EF and ToM in older children. This is of particular importance in order to understand whether both constructs remain intertwined in the course of development. It may well be that the link between ToM and EF is less relevant once sufficient cognitive capacities have developed, thereby making it less relevant to regulate one's own knowledge and view of the world when inferring others' mental states. Therefore, the first aim of the current study was to investigate whether the EF–ToM relationship extends to middle childhood, and especially, how different EF subcomponents (i.e., attention shifting, WM updating, inhibition) are related to ToM at this age.

Furthermore, an important and controversial question that remains unresolved concerns the causal direction of effect between EF and ToM. Perner (1998; Perner and Lang, 1999, 2000; see also Carruthers, 1996) and Russell (1996, 1997; see also Pacherie, 1997) both maintain that a functional dependency between the two constructs exists but they make opposite predictions as regards the direction of effect.

Perner (1998) claim that the ability to represent mental states on a meta-level is needed for the development of executive control, i.e., ToM enhances EF. In other words, this meta-representational account claims that children need to have a sufficiently developed understanding of their own minds before they will be able to engage in executive control. Russell's (1996, 1997) theory, on the other hand, claims the exact opposite, i.e., EF is a prerequisite for the emergence of ToM understanding. According to this view, EF is necessary in order to distance oneself from reality and move toward abstract mental states (ToM).

An explanation as to why there is little agreement on the causal relationship between EF and ToM lies in the fact that many studies have based their conclusions on correlational data. There are three types of evidence in order to progress beyond correlational studies (Miller, 2009). First, by means of longitudinal studies (if A predicts B, A must start developing before B). Second, by means of dissociation studies (if A causes B, then A should occur without B; but the opposite should not take place). Third, by means of training studies (if A is trained, what effect, if any, does it have on B?).

There is empirical support for all three types of evidence. First, although there are still relatively few studies examining the longitudinal relationship between EF and ToM performance, there has been a recent increase in the amount of research carried out (Hughes, 2011), including different ages and time spans ranging from very short intervals in so-called microgenetic studies, a method in which the process of developmental change is closely observed and analyzed trial-by-trial, (e.g., Flynn et al., 2004; Flynn, 2007) to longer intervals of up to 1 year (Carlson et al., 2004; Hughes and Ensor, 2007). These studies have mostly been conducted in the preschool or late toddlerhood years. A general finding at this early age is that stronger support is found for the proposal that early EF predicts later ToM than for the view that early ToM predicts later EF. For instance, one of the earliest studies showed that early EF performance (in particular inhibitory control) at age 4 predicted later ToM performance (1 year later), but the reverse was not true (Hughes, 1998b). In a study conducted with 2-year-olds, this pattern of early EF predicting later ToM (15 months later) persisted even when age, gender, and verbal ability were taken into account (Carlson et al., 2004). However, different ToM tasks were used at the two points of measurement, and therefore it is questionable whether the same construct was measured at each time point. This assumption is supported by the finding that the two ToM measures did not correlate over time. Yet, a longitudinal study that included three time points (time intervals ranging from 9 to 12.5 months), also found evidence for the view that earlier EF predicted later ToM in 2-, 3-, and 4-year-olds (Hughes and Ensor, 2007). Others have found similar results (Jahromi and Stifter, 2008; Müller et al., 2012).

However, other longitudinal studies do not support the view that early EF predicts later ToM. For instance, Schneider et al. (2005) in a study including three time points did not find predictive relationships between EF and ToM in either direction after controlling for age and language (Schneider et al., 2005). As noted by the authors, one reason for this finding might have been the fact that initial ToM understanding was low and continued to be so throughout the study period.

The focus of longitudinal studies conducted to date has mostly been on early development of EF and ToM from toddlerhood up to the preschool years. Studies in older children are rare despite the importance of examining whether patterns found in early life remain throughout the course of the child's development (McAlister and Peterson, 2013).

The few studies that have investigated the longitudinal relationship between EF and ToM in children beyond the preschool years show inconsistent results. A study by Farrant et al. (2012) in children aged 5 (at t1) found that early EF predicted later ToM, thus supporting Russell's (1996) theory and replicating general findings of longitudinal studies in younger children. However, a 1-year longitudinal study with 5-year-old children showed that early EF did not predict later ToM (Razza and Blair, 2009). But EF at t2 was not assessed, so the reverse direction of early ToM predicting later EF could not be tested. A further study in 4-year-old children (at t1) even found the opposite, namely that early ToM predicted later EF even when controlling for age, language, siblings and initial EF (McAlister and Peterson, 2013), thus in line with Perner's theory (Perner, 1998).

Taken together, longitudinal studies yielded mixed results on the exact nature of the causal relation between EF and ToM. If anything, there seems to be stronger evidence for Russel's theory that EF is a prerequisite for ToM than for Perner's theory that ToM supports EF (Russell, 1996, 1997; Perner, 1998).

The second type of evidence concerns dissociation studies. Perner's and Russel's theories make opposing predictions with respect to EF–ToM deficiencies. Perner's account excludes the possibility that well-developed EF occurs paired with poor ToM: ToM is seen as a prerequisite for EF, and therefore ToM deficits should lead to impaired EF. But what is in line with his theory is the reverse pattern of well-developed ToM paired with EF deficits because intact ToM is necessary for EF, but not sufficient in its own right. On the other hand, Russel's theory does not allow for deficits in EF to be paired with adequate ToM because his theory proposes that impaired EF leads to impaired ToM. His theory does permit the pattern of well-developed EF paired with deficits in ToM because the ToM impairment could have been caused by other factors, e.g., language, in particular inner speech (Pellicano, 2007).

Several dissociation studies indicate that EF is not required for the development of ToM, thus contradicting Russel's theory. For example, a study conducted by Tager-Flusberg et al. (1997) and reanalyzed by Perner and Lang (2000) showed that children with Williams syndrome and Prader-Willi syndrome were impaired on EF tasks but mastered ToM tasks well. However, due to the small sample size ($N = 6$), results must be treated with caution. Studies on children diagnosed with ADHD showed a similar pattern, i.e., reasonably well-developed ToM skills paired with poor EF (Charman et al., 2001; Perner et al., 2002). However, it may well be that children in these studies mastered the ToM tasks in atypical ways, for instance by applying simple behavior rules (e.g., "doesn't see, doesn't know"; Garnham and Ruffman, 2001). If so, the relevance of EF for ToM would not be challenged by these results.

Other dissociation studies have found evidence contradicting Perner's theory. Although autism typically involves low EF paired with low ToM (e.g., Hill, 2004), a study in $5\frac{1}{2}$ year-old children with autism revealed a dissociation in exactly the opposite direction: a high level of EF with impaired ToM (Pellicano, 2007). Similarly, a cross-cultural study comparing U.S. and Chinese children aged about 4 years, controlled for age and verbal ability, found that although Chinese children had good executive control skills, their ToM understanding was poor (Sabbagh et al., 2006). Thus, evidence from dissociation studies revealed support for both Perner's and Russel's theories.

The third and final type of empirical support comes from training studies. For instance, a very recent study revealed the importance of EF for the improvement in the development of ToM. Preschool-aged children (aged 3 years 8 months) were given a battery of false-belief and EF tasks. Results showed that individual differences in initial EF predicted the degree to which children's advances in ToM improved through ToM training, the relevant control variables being partialled out (Benson et al., 2013). However, a training study by Kloo and Perner (2003) did not support the view that early EF predicts later ToM to the same extent. They examined ToM via false-belief tasks, and attention shifting via the dimensional change card sort task (DCCS; Frye et al., 1995). Children (3- to 4-year-olds) were trained in these tasks about once a week over a period of $\sim$ weeks. Transfer of training took place in both directions: training on the DCCS task improved performance on the false-belief task, and false-belief training produced improvements on the DCCS task. This finding is in support of the idea of functional dependency between EF and ToM. However, it is interesting to note that training on the false-belief task did not lead to an improvement in post-training false-belief performance, which makes the interpretation of these findings problematical.

To sum up the three types of evidence, the general finding (particularly from longitudinal studies) in children younger than 4 years is that early EF predicts later ToM, and not vice versa. However, studies in older children have revealed inconsistent results. More research is needed to clarify the causal direction between EF and ToM in older children. In particular, possible differential relations between EF subcomponents and ToM are understudied to date. Therefore, the second aim of the present study was to examine whether each of the three EF subcomponents (attention shifting, WM updating, inhibition) at t1 predicted ToM at t2 (1 year later), or vice versa, in elementary school-aged children.

The current study investigated three subcomponents of EF (i.e., inhibition, attention shifting, WM updating) as well as ToM in a large sample of 6- to 12-year-old children. Moreover, we examined relations between EF and ToM controlled for age, gender and fluid intelligence. All tasks were administered at two measurement points which were about 1 year apart. Based on the evidence reviewed so far, we hypothesized first, the EF–ToM relationship to pertain in middle childhood. Thus, positive correlations should occur between EF and ToM performance in our sample. More specifically, we hypothesized that the three EF subcomponents and ToM at t1 are as strongly correlated as 1 year later at t2. Second, we expected to find relationships leading from EF subcomponents-t1 to ToM-t2, and not vice versa.

To our knowledge, this is the first study that addresses these issues in a representative sample of children in middle childhood.

MATERIALS AND METHODS

PARTICIPANTS

At t1 (in 2012), the sample consisted of $N = 1{,}657$ children (52% girls) aged between 6 and 11 years ($M = 8.3$ years, $SD = 0.95$). Time point 2 (t2; in 2013) took place 1 year later ($N = 1{,}619$) and children's age ranged from 7 to 12 years ($M = 9.1$ years, $SD = 0.92$).

Participants were recruited from 33 primary schools from the federal state of Brandenburg, Germany. To establish a representative sample, schools were preselected so that participants coming from different rural and urban areas and socio-economic backgrounds were included. Before the study began, approval of all procedures was granted by the Research Ethics Committee at the University of Potsdam and the Ministry of Education, Youth, and Sports of the Federal State of Brandenburg. For each child, informed consent was obtained from the parent/primary caregiver. As a reward the children received a voucher for the cinema.

MATERIAL

EF tasks

The EF subcomponent attention shifting was assessed using the Cognitive Attention shifting task (Röthlisberger et al., 2010; adapted from Zimmermann et al., 2002). Children were presented with a single-colored fish and a multi-colored fish appearing simultaneously on the left- and right-hand side of a computer screen. Participants were told to "feed" each kind of fish, alternating between the two, by pressing one of two keys on a QWERTZ keyboard (i.e., the X-key for the fish on the left, the M-key for the fish on the right). Across several trials, the side on which the two kinds of fish appeared changed randomly, requiring the children to remember their response of the previous trial in order to maintain the requirement of alternating feeding. The task consisted of a total of 46 trials (interstimulus intervals ranged from 300 to 700 ms) that were separated by a short break during which positive feedback was given. The dependent variable was the number of correct responses for the 22 switch trials. Switch trials were those answers that required children to change their response pattern (i.e., from pressing left/right to left/left or right/right).

The EF subcomponent WM updating was assessed using the Digit-Span Backward task (Petermann and Petermann, 2007). Participants were told a sequence of digits which they had to verbally repeat in the reverse order. The first sequence was two digits long. There were two sequences of equal length in each trial. Within each trial, at least one sequence had to be answered correctly in order to proceed to the next trial in which the sequence was lengthened by one digit. The dependent variable was the total number of sequences that had been remembered correctly.

The EF subcomponent inhibition was assessed by the Fruit Stroop task (Archibald and Kerns, 1999; adapted by Röthlisberger et al., 2010). The task consisted of four trials. For each trial, a page depicting 25 stimuli was presented to the child with the instruction to name the colors of the items as quickly as possible. Page 1 showed colored rectangles (blue, red, green, yellow). Page 2 depicted fruits or vegetables in their typical colors (banana – yellow, lettuce – green, strawberry – red, plum – blue). Page 3 showed the same fruits but these were colored gray. Page 4 showed the same fruits and vegetables, but all colored incorrectly. For pages 3 and 4, the children had to name the color that the fruit and vegetables should have had (i.e., banana – yellow, lettuce – green, etc.). The time (in seconds) required for naming the colors of all 25 items per page was measured. As dependent variable, an interference score was generated: time p.4 – [(time p.1 × time p.3)/(time p.1 + time p.3); Archibald and Kerns, 1999)]. Scoring high on this task is an indication of a lower ability to successfully inhibit the prepotent response of naming the items' actual colors on page 4.

ToM task

Theory of mind ability was assessed by a cartoon task developed by Völlm et al. (2006) for adults and adapted by Sebastian et al. (2012) for adolescents. The cartoon scenarios were presented on a computer screen. Each story consisted of five pictures with black-and-white line drawings depicting two individuals in order to control for social content (**Figure 1**). The first three story frames appeared consecutively, followed by two pictures shown simultaneously that displayed possible endings of the story. Children were instructed to choose the correct ending by pressing the X-key on a QWERTZ keyboard for the left-hand side, and the M-key for the right-hand side picture. In order to give a correct answer, children had to infer the mental state of one protagonist and the appropriate response by the other protagonist. Interstimulus intervals ranged from 1,000 to 3,000 ms. The order of the cartoons and the location of the correct ending were randomized across participants. Correct responses for each story were coded as 1, incorrect responses as 0.

The original task by Sebastian et al. (2012) consisted of 30 cartoons, 10 in each of three conditions: cognitive ToM, affective ToM, and a physical control condition. The physical control condition was used as baseline in order to determine neuronal activity in an functional magnetic resonance imaging (fMRI) setting. Because the current study did not use fMRI, and a baseline condition was not relevant, the physical condition was excluded. The remaining 20 cartoons (10 cognitive, 10 affective) were pretested for clarity, difficulty, and timings in children who were of similar age to those in our study and did not participate in the present study. Twelve cartoons (six cognitive ToM, six affective ToM) of mid-range difficulty were selected in order to limit the time demands of the task and to ensure sufficient variability in the data.

Each child was presented with six cognitive and six affective ToM stories. To choose the correct ending for the cognitive scenarios, children had to infer the appropriate behavior of one protagonist (e.g., helping) given the inferred intentions, desires, or beliefs of the other protagonist (e.g., attaining an action goal; **Figure 1A**). For the affective scenarios, participants needed to infer the appropriate response (e.g., consoling) of one protagonist regarding the emotional state (e.g., fear, sadness) of the other (**Figure 1B**). Performance on cognitive and affective trials was highly associated in our study (at t1: $r = 0.83$, $p \leq 0.001$; at t2: $r = 0.94$, $p \leq 0.001$). Thus, for reasons of simplicity and in line with the focus of the current paper, a model that did not differentiate between cognitive and affective aspects was chosen for further analyses. This one-factor model fitted the empirical data well at both time points, t1: $\chi^2(9) = 17.11$, $p < 0.05$, $CFI = 0.99$, $RMSEA = 0.024$, $WRMR = 0.765$, t2: $\chi^2(9) = 14.85$, $p = 0.095$, $CFI = 0.988$, $RMSEA = 0.020$, $WRMR = 0.650$. However, two items were excluded due to factor loadings falling under a general cutoff value (0.40) for the inclusion into one factor (Stevens, 2001).

In addition, the final version of the task was validated in a sample of 7.5- to 10-year-old children ($N = 62$, M age $= 8.23$,

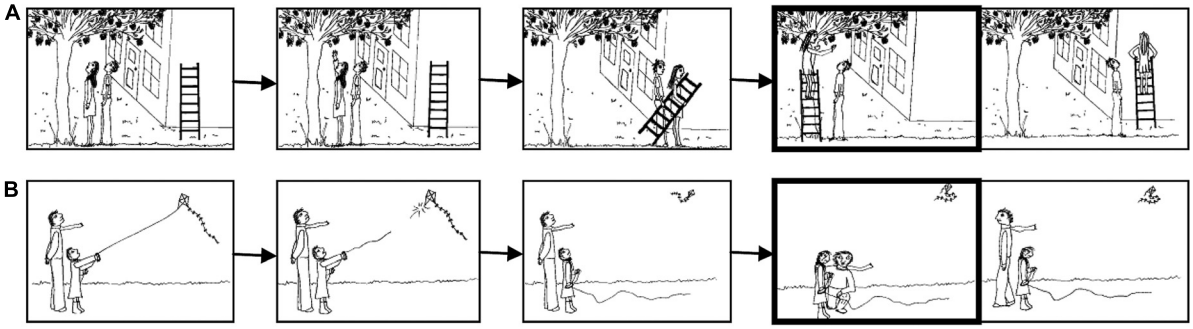


FIGURE 1 | Five pictures of one cartoon story for cognitive ToM (A) and affective ToM (B). The first three pictures appeared on a computer screen consecutively, the last two pictures appeared simultaneously. Children were instructed to select the correct ending by pressing a key. For illustration, the frame of the correct ending is marked black.

SD = 0.59, 35.5% girls) in order to ensure its association with standard ToM measures.

The standard ToM measures employed were two second-order false-belief tasks (Perner and Wimmer, 1985; Hollebrandse et al., 2012), a German version of the Extended Theory-of-Mind Scale (Henning et al., 2012) and German translations of the Strange Stories Test (Happé, 1994). Results showed that the total score of the standard ToM measures was positively correlated with the ToM cartoons total score (affective and cognitive cartoons combined) after controlling for language ($r = 0.29$, $p = 0.025$) or fluid intelligence ($r = 0.32$, $p = 0.047$).

Fluid intelligence

To measure fluid intelligence the Number-Symbol Test of the German version of the Wechsler Intelligence Scale for Children was used (Petermann and Petermann, 2007). The child was instructed to redraw symbols (e.g., a half moon) that were paired with either simple figures (e.g., a cross with a circle inside; Version A for 6- to 7-year-olds) or digits (1–9; Version B for 8- to 16-year-olds) as quickly as possible within 120 s. For both versions A and B, the dependent variable was the amount of correct symbols allocated within 120 s. For version A, extra points could be achieved if participants finished the task before the 120 s were over.

PROCEDURE

At both t1 and t2, children were assessed for two 50-min sessions spaced 1 week apart. Assessments were part of a larger battery of tasks that were conducted separately with each child by a specifically trained PhD student or research assistant in a quiet room either in a school setting or at home. The order of the larger battery of tasks was counterbalanced across participants (blocks of ABCD/BADC) but the order did not show any effect when subsequently analyzed.

DATA ANALYSIS

All analyses were run using Mplus 7.11 (Muthén and Muthén, 1998–2012). The rate of missing data was low on all variables at t1 ($\leq$1.8) and at t2 ($\leq$7.2). Assuming data to be missing at random, all missing values were accounted for by full information maximum likelihood (FIML) estimation.

Research questions were answered by means of structural equation modeling (SEM). As mentioned above, model fit was considered acceptable only if absolute fit indices fulfilled the following criteria: $CFI \geq 0.95$, $RMSEA \leq 0.08$, $WRMRV \leq 1.0$ (Yu, 2002; Geiser, 2010).

The cartoon stories were entered in the analyses as categorical indicators (1 for choosing the correct ending, 0 for an incorrect choice). As the χ^2 depends on sample size and is overly sensitive to deviations from perfect fit in large samples (Schermelleh-Engel et al., 2003).

In order to answer the first research question (the extension of the EF–ToM relationship to middle childhood) the concurrent correlations between each EF subcomponent and ToM at t1 and t2 were examined (EF-t1 – ToM-t1 and EF-t2 – ToM-t2). In addition, the correlation coefficients at t1 and t2 were statistically compared in order to determine whether the EF–ToM association was equally strong at each time point. This procedure was followed for each of the three EF subcomponents in order to detect possible differential relations.

To answer the second research question (the longitudinal relations between EF and ToM over a 1-year period), three cross-lagged models were fit to the data, describing the assumed interrelations between each EF subcomponent and ToM over time. By controlling for initial levels of the target variable, cross-lagged models ensure that the association of one variable as developmental precursor of another variable is examined, rather than concurrent associations (McAlister and Peterson, 2013). Subsequently, regression coefficients for the cross-lagged paths were compared in order to evaluate for which direction the association was stronger, i.e., ToM-t1 to EF-t2 or EF-t1 to ToM-t2.

In all analyses, ToM was entered as latent and the three EF subcomponents (attention shifting, WM updating and inhibition) as manifest variables. Moreover, fluid intelligence, age and gender which are known to be related to EF and ToM were included as manifest control variables.

RESULTS

DESCRIPTIVES

Descriptive statistics of the assessed variables are shown in **Table 1**. At both measurement time points, medium scores were achieved

Table 1 | Descriptive statistics of assessed variables (EF subcomponents attention shifting, WM updating, inhibition, and ToM cartoon task) for the two measurement time points (t1 and t2) and mean comparison results.

	t1			t2			*t*
	M	***SD***	**Min.–max.**	***M***	***SD***	**Min.–max.**	
Attention shifting	15.6	4.7	0–22	18.1	3.9	0–22	−24.2***
WM updating	6.1	1.6	0–16	6.6	1.5	0–16	−11.1***
Inhibition[a]	24.9	8.8	0–89[b]	20.6	6.9	0–66[b]	23.2***
ToM[c]	8.5	1.7	0–10	9.2	1.3	1–10	−14.5***
Fluid intelligence[d]	51.4	9.2	27–80[b]	/	/	/	/

ToM, theory of mind; [a] *interference measure (negatively polarized);* [b] *min and/or max values are theoretically infinite, thus table values are sample-specific;* [c] *average number of correct trials;* [d] *only the t1 measurement of fluid intelligence was included in the analysis; ***p ≤ 0.001.*

on WM updating, inhibition and fluid intelligence, medium to high scores on attention shifting and high scores on ToM. Intercorrelations between the three EF subcomponents ranged from 0.27 to 0.35 (all *p*s ≤ 0.001) at t1 and from 0.28 to 0.33 at t2 (all *p*s ≤ 0.001). On average, participants improved in the EF and ToM tasks from t1 to t2, indicating a significant developmental change in those abilities within a year.

RESEARCH QUESTION 1: CONCURRENT ASSOCIATIONS BETWEEN EF SUBCOMPONENTS AND ToM

The first research question concerned the concurrent correlations at both time points controlled for age, gender, and fluid intelligence, that is, the association between EF subcomponents and ToM at t1 and t2 (see **Table 2**). At both time points, correlations were small but significant (ranging from 0.10 to 0.20; Cohen, 1988) indicating that EF subcomponents and ToM were associated in 6–11 year-olds and in 7–12 year-olds. Testing the strength of the concurrent path coefficients against each other showed that for each EF subcomponent the difference of t1 and t2 concurrent correlations with ToM was not significant (attention shifting: $\Delta\chi^2$ (1) = 0.041, $p = 0.84$; WM updating: $\Delta\chi^2$ (1) = 3.285, $p = 0.07$; inhibition: $\Delta\chi^2$ (1) = 0.037, $p = 0.85$).

Table 2 | Concurrent correlations between EF subcomponents and ToM at t1 and t2, respectively, controlled for age, gender, and fluid intelligence.

	ToM-t1		ToM-t2
Attention shifting-t1	0.19**	Attention shifting-t2	0.20**
WM updating-t1	0.13**	WM updating-t2	0.20**
Inhibition-t1[a]	−0.08*	Inhibition-t2	−0.10**

[a] *Interference measure (negatively polarized); **p ≤ 0.01, *p ≤ 0.05, two-tailed.*

RESEARCH QUESTION 2: RECIPROCAL INFLUENCES BETWEEN EF SUBCOMPONENTS AND ToM ACROSS 1 YEAR

Figure 2 shows the cross-lagged models for the interrelation between each EF subcomponent and ToM over time controlled for initial levels of the outcome variable at t1 as well as for the covariates age, gender, and fluid intelligence.

The model for attention shifting (**Figure 2A**) fitted the data well, χ^2 (259) = 321.13, $p = 0.005$, *CFI* = 0.98, *RMSEA* = 0.012, *WRMR* = 0.90. Both cross-lagged path coefficients (attention shifting-t1 to ToM-t2 and ToM-t1 to attention shifting-t2) reached significance. Despite the high autocorrelations of attention shifting and ToM from t1 to t2 (**Figure 2A**), and after controlling for age, gender, and fluid intelligence, the cross-lagged paths still revealed a small but significant association of EF on ToM, and vice versa, across the 1-year period. Testing the strength of the cross-lagged path coefficients against each other showed a significant difference, $\Delta\chi^2$ (1) = 7.999, $p < 0.01$, with a stronger association leading from attention shifting-t1 to ToM-t2 than in the opposite direction.

The model for WM updating (**Figure 2B**) also fitted the data well, χ^2 (259) = 328.75, $p = 0.002$, *CFI* = 0.98, *RMSEA* = 0.013, *WRMR* = 0.91. Again, both cross-lagged path coefficients (WM-t1 to ToM-t2 and ToM-t1 to WM-t2) revealed a small, but significant impact over and above the high autocorrelations over time and possible effects of the control variables. The difference between the cross-lagged path coefficients was marginally significant, $\Delta\chi^2$ (1) = 3.42, $p = 0.064$, with a stronger association leading from WM-t1 to ToM-t2 than in the opposite direction.

The model for inhibition (**Figure 2C**) also fitted the data well, χ^2 (259) = 316.28, $p = 0.009$, *CFI* = 0.98, *RMSEA* = 0.012, *WRMR* = 0.90. However, for this model neither of the cross-lagged path coefficients reached significance. Thus, for the EF subcomponent inhibition no significant cross-lagged relationships with ToM over time were found.

DISCUSSION

The present study pursued two objectives. First, we examined whether the relationship between EF and ToM pertains in middle childhood (6–12 years). Results showed small but significant concurrent correlations in the expected directions between all studied EF subcomponents (attention shifting, WM updating and inhibition) and ToM. In line with previous research (Perner et al., 2002; Yang et al., 2009; Calderon et al., 2010), better abilities in executive control of thought or action were related to better

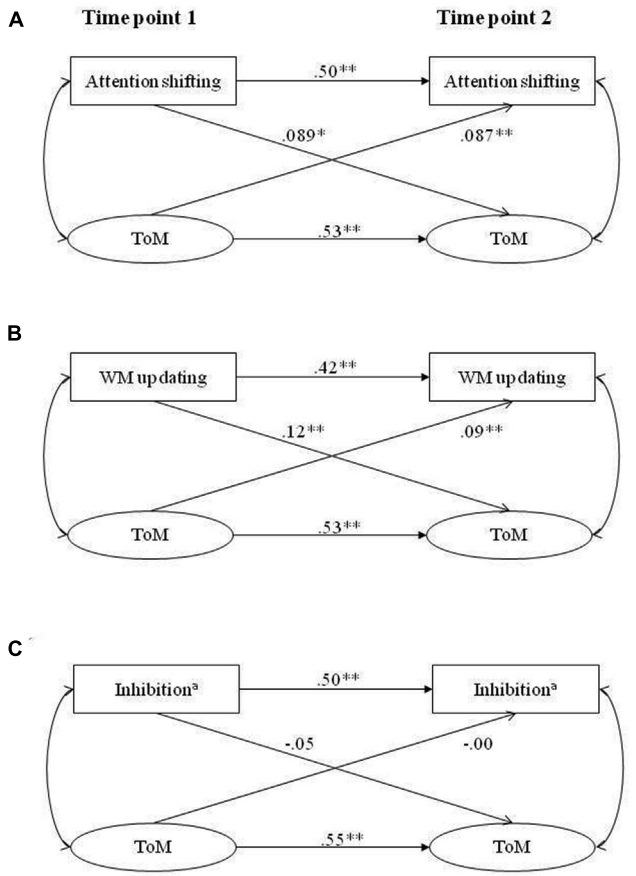


FIGURE 2 | Cross-lagged models for the relations between ToM and the EF subcomponents attention shifting (A), WM updating (B), and inhibition (C), at the two measurement points, controlling for age, gender, fluid intelligence, as well as for the earlier levels of the target variables. [a]Interference measure (negatively polarized); $^{**}p \leq 0.01$, $^{*}p \leq 0.05$, two-tailed.

understanding of others' mental states at t1 (6–11 years) and t2 (7–12 years). Second, we explored whether each EF subcomponent at t1 predicted ToM at t2, or vice versa, over a 1-year period. Here, we used a cross-lagged model, again controlling for age, gender and fluid intelligence, as well as for the earlier levels of the target variables. Results showed small, but significant bidirectional longitudinal relationships with ToM for two EF subcomponents, WM updating and attention shifting. Examining the strength of the associations showed that for both EF subcomponents, the relationship between early EF and later ToM was stronger than the relationship between early ToM and later EF. Bearing in mind that effect sizes were small this corresponds with research on longitudinal studies between EF and ToM in preschool-age children (e.g., Hughes, 1998b) and illustrates the pertaining relevance of the ability to switch between different task demands and the ability to temporarily hold information in mind while processing it for developing ToM abilities. For the subcomponent inhibition, however, no reciprocal relationships were found over time. Thus, discriminating between EF subcomponents seems to be important for the study of EF–ToM development, because on the one hand, the single EF subcomponents may follow different developmental courses and on the other hand, understanding the mental states of others at various levels (e.g., first- or second-order perspective) may put different demands on EF subcomponents.

THE RELATIONSHIP BETWEEN THE THREE EF SUBCOMPONENTS AND ToM IN MIDDLE CHILDHOOD

The relationship between EF subcomponents and ToM is well documented in the preschool years (e.g., Frye et al., 1995; Hughes, 1998a; Carlson and Moses, 2001). However, still in need of clarification is the question how exactly this relationship extends to middle childhood for each of the EF subcomponents. To date, there are few studies in children beyond the preschool age, and, as we will discuss below, those conducted are inconclusive. Results of the present study indicate that the relationship between all three EF subcomponents (attention shifting, WM updating, inhibition) and ToM pertain in middle childhood with small, but significant correlations (age, gender, and fluid intelligence partialled out).

In a sample of similar-aged children (8.5-year-olds) Charman et al. (2001) also found a correlation between EF (GoNoGo error score) and ToM (Happé stories correct mental score) for their typically developing control group ($r = -0.43$, $p < 0.01$). However, as soon as age and intelligence were partialled out, their results fell below significance ($r = -0.38$, $p = 0.10$). This may be owing to the low statistical power due to the relatively small size of their control group ($N = 22$). The diverging results to the present study may also come from the use of different measures. Charman et al. (2001) assessed ToM by means of Happé's Strange Stories (Happé, 1994), a measure of higher-order ToM ability. The Strange Stories task consists of naturalistic, short vignettes which are read to the child by the experimenter. The task makes strong demands on verbal ability. In contrast, in the present study, a ToM cartoon task was used. It is a non-verbal task that was originally developed for fMRI studies (Völlm et al., 2006). Each cartoon story is presented on a computer screen, and the correct answer is chosen by pressing a key. Because of the lower verbal task demands of the ToM cartoon test, children in the present study were probably not impaired in expressing their ToM ability.

Turning to EF, Charman et al. (2001) examined different subcomponents (planning and behavioral inhibition) compared to our study. Generally, the fact that different studies involve various aspects of EF makes comparisons between studies difficult. This inconsistency might be due to the fact that EF is an ill-defined construct that has been described as an umbrella term for a large array of different processes involved in self-control (Kerr and Zelazo, 2004). The EF definition applied in the current study follows Miyake et al.'s (2000) division into three overlapping but distinct subcomponents – attention shifting, WM updating, and inhibition. This approach has proved promising and has been adopted in studies on the relations between ToM and EF in preschool children.

Another problem when comparing different studies on EF–ToM relations is that the available tasks do not test only one EF subcomponent, but engage overlapping EF aspects to varying degrees. The planning tasks employed by Charman et al. (2001)

displays this *task impurity* to a rather high degree because more than one EF subcomponent was needed to meet the relatively complex task requirements (Miyake et al., 2000). Because the present study used simpler tasks, which mainly required one EF subcomponent, children could probably express their EF capacities in an optimal way. This may be another reason why we, unlike Charman et al. (2001) still found small but significant EF–ToM relations in 6- to 12-year-olds after controlling for age, gender, and intelligence.

Our results are supported by a study in slightly younger children ($4^1/_2$–$6^1/_2$ year-olds) where strong relationships between several EF and ToM tasks were found even when controlling for age and IQ and despite the relatively small sample size ($N = 22$) of their control group (Perner et al., 2002). These similar findings can possibly be attributed to the fact that similar measures were implemented. Perner et al. (2002) used a second-order false-belief task (based on the material by Perner and Wimmer, 1985). We argue that at least 6 out of 10 of our ToM cartoon stories require second-order reasoning. Because the cartoons were originally used in a study with a different focus - differentiating ToM into cognitive and affective aspects - Sebastian et al. (2012) did not address the issue of first- or second-order ToM reasoning. What they did maintain was that the affective condition requires cognitive ToM. Yet, in the six affective ToM cartoon tasks, in order to give a correct response, the participant has to understand the first protagonist's belief about the second protagonist's mental state (e.g., the adult *believes/thinks* the child is sad because the child does not *want* the kite to fly away; i.e., second-order ToM). In the four cognitive ToM cartoon tasks, however, in order to choose the correct ending, the participant has to infer the desire of only one of the protagonists (e.g., the girl places the ladder against the tree because she *wants* an apple; i.e., first-order ToM). The second protagonist merely accompanies the first protagonist.

We also tested possible differences in the strength of the concurrent relationship between EF subcomponents and ToM at both time points. We had no reason to hypothesize that there would be any change in the EF–ToM relations in the space of 1 year in 6 to 12-year-olds. The present results showed precisely this, suggesting that all three EF subcomponents and ToM remain related, with not much change in the strength of the relations across a 1-year-period, in a representative sample in middle childhood.

THE RECIPROCAL RELATIONSHIP BETWEEN EF SUBCOMPONENTS AND ToM

The second aim of the present study explored the longitudinal bidirectional relationship between the three EF subcomponents and ToM over a 1-year period. In a cross-lagged model, we again controlled for age, gender, and fluid intelligence, and also partialled out earlier EF or ToM performance. Our findings revealed small, but significant reciprocal relations with ToM for two EF subcomponents, attention shifting and WM updating, but no relations for inhibition. The lacking longitudinal relations between inhibition and ToM may indicate that this EF subcomponent is of less importance for ToM in middle childhood as compared to the preschool age, particularly over time. The different results for the three EF subcomponents underlines the necessity of examining EF development divided into its specific subcomponents. Interestingly, for both attention shifting and WM updating, the relationship of early EF predicting later ToM was stronger than the relationship of early ToM predicting later EF (the difference for WM updating was marginally significant). Although our effect sizes were small, this finding supports Russell's (1996) theory that EF precedes ToM development and is in line with results of previous longitudinal studies on EF–ToM relationship in preschool-age children.

One of the earliest such studies found that in 3- to 4-year-old children (mean age: 3 years and 11 months) early EF performance predicted later ToM (13 months later) even when controlling for age, verbal ability, and initial ToM (Hughes, 1998b). The reverse direction was not found, which has been interpreted as evidence for Russel's theory. However, only one of the EF measures, a detour-reaching box (Hughes and Russell, 1993), a measure of inhibitory control, showed an independent predictive relationship with later ToM in Hughes's study. The other EF subcomponents (WM, planning, attention shifting) did not. Interestingly, in the present study with older children, the only two EF subcomponents that showed small but significant relationships with ToM over time were attention shifting and WM updating, but not inhibition. A possible explanation for this finding may be that preschool-age children are still in the course of developing their inhibitory skills and thus rely on these more heavily. This is reflected in medium to high longitudinal correlations between inhibition and ToM. In addition, it may also be that first-order false-belief tasks (as used by Hughes) make more demands on inhibitory skills compared to second-order ToM items (as mainly used in the present study). In first-order false-belief tasks, the child's own knowledge about reality has to be inhibited in order to give a correct answer. To solve second-order ToM tasks, inhibition may play less of an important role, because the focus lies more on being able to switch flexibly between the different mindsets of the first and second protagonist (Miller, 2009). Also, WM updating is needed to keep track of the different perspectives and bringing all the relevant information together. The exact nature of the relationship between different EF subcomponents and ToM across childhood requires further research.

Just like Hughes's (1998b) study, Carlson et al.'s (2004) longitudinal study in a younger group of children (2-year-olds) showed similar asymmetrical relationships: early EF predicted later ToM (15 months later) even when controlling for age, gender, and verbal ability. However, as mentioned in the introduction, Carlson et al.'s (2004) results must be interpreted with caution as ToM was assessed with different measures at the two time points and did not correlate over time.

Another longitudinal study by Hughes and Ensor (2007) examined the predictive relationship between EF and ToM at three time points in 2-, 3-, and 4-year-olds. Results showed only limited support for Perner (1998), Perner and Lang's (1999) theory – that early ToM predicts later EF – but stronger support for Russell's (1996) theory – that early EF predicts later ToM. One advantage of their study was that they included participants from a variety of social backgrounds. This issue has been neglected in previous research despite the fact that socioeconomic status (SES) is known to predict cognitive abilities (Bradley and Corwyn, 2002) and may have an impact on the EF–ToM relationship (Hughes and

Ensor, 2007). Although only little is known about the impact of SES on EF or ToM development in middle childhood, the present study established a representative sample by preselecting schools from different socio-economic backgrounds in both rural and urban communities. Therefore, the present results are probably not affected by possible SES effects.

Turning now to our findings of reciprocity, although the present study showed stronger relationships leading from early attention shifting and WM updating to later ToM, one cannot dismiss the fact that the relationship was bidirectional. That is to say, there were indeed small, but significant relations in the opposite direction. We did not expect this finding because previous longitudinal studies almost consistently showed a unidirectional association in which early EF predicted later ToM and not vice versa (Hughes, 1998b; Carlson et al., 2004; Flynn et al., 2004; Jahromi and Stifter, 2008; Müller et al., 2012). However, in line with our findings, Hughes and Ensor (2007) also found partial support for early ToM – later EF, but stronger evidence for early EF – later ToM. Likewise, McAlister and Peterson (2013) found that ToM at 3 years 3 months to 5 years 6 months predicted EF at about 1 year later, but they did not find a significant relation in the opposite direction. Moreover, Kloo and Perner's (2003) training study in 3- to 4-year-olds suggested a bidirectional relationship between EF and ToM because transfer of training took place in both directions. They took their results to support the idea of a functional dependence between the two constructs in that "understanding the mind presupposes a certain degree of executive control, and EF presupposes a certain level of insight into the mind" (Kloo and Perner, 2003, p.1836). It has been suggested elsewhere that the EF–ToM relationship can be interpreted in reciprocal terms with one construct complementing the other (Putko, 2009). However, as noted by Kloo and Perner (2003), their study cannot clarify the causes and processes involved that are responsible for this relationship. Both constructs may be related in an indirect way, that is, individuals with well-developed EF may be better equipped to function well in social groups and this then encourages improvements in ToM (Hughes, 2001). Further studies are needed to shed more light on the relationship between EF and ToM in reciprocal terms and on the exact interplay of the two constructs, especially for the age range beyond the preschool years.

In sum, although a few studies have shown that early ToM predicts later EF, the majority of longitudinal studies conducted so far reveal more evidence for the view that early EF predicts later ToM. However, the existing longitudinal research has almost exclusively focused on the preschool or late toddlerhood years. To our knowledge, the current study is the first to find longitudinal support suggesting that early EF subcomponents predict later ToM in a representative sample in middle childhood. This is an interesting finding as it suggests that although sufficient cognitive capacities have developed, a higher level of EF (WM updating and attention shifting) continues to be associated with a higher level of ToM. Thus, especially the ability to switch between different task demands and the ability to temporarily hold information in mind while processing it seem to be important for understanding mental states in middle childhood. The importance of these two EF subcomponents is reflected in second-order false-belief tasks, the most commonly used ToM task in older children. In order to make a correct inference on this task, the different perspectives of the mindsets of the two protagonists need to be taken into account (attention shifting) and, in addition, all the pieces of information need to be actively held in mind and updated (WM updating). Thus it appears that while children progress through childhood and their social contexts become increasingly complex (e.g., school entry in which friendships and relationships to peers and teachers are formed), attention shifting and WM updating are needed for the children to make sense of and function within their social surroundings.

LIMITATIONS AND OUTLOOK

A problem which has been discussed in the literature is the possible difficulty that children may have to express their existing ToM abilities due to the EF requirements inherent in the ToM task (Carlson and Moses, 2001; Moses, 2001). Referring to the instruments in our study, it is possible that our ToM cartoon requires EF demands which may explain the concurrent as well as the longitudinal correlations between the two constructs. However, several arguments speak against this. Many children completed our ToM cartoon task successfully, which implies that the EF demands were not overly high. In addition, it may be that our ToM cartoon task measures WM capacity, producing a task impurity problem (Miyake et al., 2000). For each cartoon, three pictures were shown consecutively, which had to be remembered by the child in order to choose the correct ending. But because the pictures were presented in quick succession, they resembled a cartoon film strip with an easy script, not making overly high demands on WM capacity. However, even if the latter were to be the case, children were not required to mentally process or re-arrange the pictures in any way. Therefore, these potential task demands cannot explain the relations with updating of WM which was one of our EF subcomponent.

Another argument on the same lines is that the relation between attention shifting and ToM may result from the fact that both tasks require shifting between pressing the left and right key. However, the ToM cartoons did not call for shifting between different answer sets or for abandoning an acquired rule of alternating between pressing the left and right key (as the attention shifting task did). In the 10 ToM cartoons, children just had to decide which of the two presented pictures displayed the correct ending, and they had ample time to press the appropriate key. The high rate of correct responses for the ToM cartoons indicates that this task was rather easy for the children. Therefore, we take the relations between attention shifting and ToM as evidence that our ToM tasks required attention shifting with respect to inferring the mental states of the two displayed protagonists (rather than with respect to the task design).

Second, further studies on longitudinal relations between EF subcomponents and ToM in middle childhood should include more than two time points. This would not only show how the EF–ToM relationship develops over a longer period, but also allow an investigation of moderating or mediating factors.

Furthermore, although our study yielded innovative findings, it would be interesting to use more than one ToM and more than one task for each EF subcomponent. Using more measures would

no doubt increase reliability, and it would shed further light on relations between EF subcomponents and different aspects of ToM that emerge in middle childhood (e.g., second-order false-belief, irony, contrary emotion; Happé, 1994).

In conclusion, the current study suggests that the relationship between three EF subcomponents (attention shifting, WM updating, inhibition) and ToM pertains in a representative sample in middle childhood. Partial evidence was provided for the assumption that early ToM predicts later EF (Carruthers, 1996; Perner, 1998; Perner and Lang, 1999, 2000) but there was stronger support for the proposal that early EF (attention shifting, WM updating) predicts later ToM (Russell, 1996). In addition, our results suggest a reciprocity in the EF–ToM relationship over a 1-year-period. Future studies are needed to shed more light on the precise interplay of the two constructs, especially with respect to subcomponents of both EF and ToM, in the course of childhood development.

ACKNOWLEDGMENTS

This research was supported by a grant from the *Deutsche Forschungsgemeinschaft* (DFG; GRK 1668/1). The authors wish to thank Sebastian Grümer and Julia Tetzner for their helpful methodological suggestions and Carolin Russek for her considerable help in collecting the data.

REFERENCES

Archibald, S. J., and Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychol.* 5, 115–129. doi: 10.1076/chin.5.2.115.3167

Benson, J. E., Sabbagh, M. A., Carlson, S. M., and Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Dev. Psychol.* 49, 1615–1627. doi: 10.1037/a0031056

Bradley, R., and Corwyn, R. (2002). Socioeconomic status and child development. *Annu. Rev. Psychol.* 53, 371–399. doi: 10.1146/annurev.psych.53.100901.135233

Calderon, J., Bonnet, D., Courtin, C., Concordet, S., Plumet, M.-H., and Angeard, N. (2010). Executive function and theory of mind in school-aged children after neonatal corrective cardiac surgery for transposition of the great arteries. *Dev. Med. Child Neurol.* 52, 1139–1144. doi: 10.1111/j.1469-8749.2010.03735.x

Carlson, S. M., Mandell, D. J., and Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Dev. Psychol.* 40, 1105–1122. doi: 10.1037/0012-1649.40.6.1105

Carlson, S. M., and Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333

Carruthers, P. (1996). "Autism as mindblindness: an elaboration and partial defense," in *Theories of Theories of Mind*, eds P. Carruthers and P. K. Smith (Cambridge: Cambridge University Press), 257–273.

Charman, T., Carroll, F., and Sturge, C. (2001). Theory of mind, executive function and social competence in boys with ADHD. *J. Emot. Behav. Disord.* 6, 31–49. doi: 10.1080/13632750100507654

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences,* 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Davis, H. L., and Pratt, C. (1995). The development of children's theory of mind: the working memory explanation. *Aust. J. Psychol.* 47, 25–31. doi: 10.1080/00049539508258765

Doherty, M. J. (2009). *Theory of Mind: How Children Understand Others' Thoughts and Feelings.* Hove: Psychology Press.

Farrant, B. M., Maybery, M. T., and Fletcher, J. (2012). Language, cognitive flexibility, and explicit false belief understanding: longitudinal analysis in typical development and specific language impairment. *Child Dev.* 83, 223–235. doi: 10.1111/j.1467-8624.2011.01681.x

Flynn, E. (2007). The role of inhibitory control in false belief understanding. *Infant Child Dev.* 16, 53–69. doi: 10.1002/icd.500

Flynn, E., O'Malley, C., and Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Dev. Sci.* 7, 103–115. doi: 10.1111/j.1467-7687.2004.00326.x

Frith, C. D., and Frith, U. (1999). Interacting minds-a biological basis. *Science* 286, 1692–1695. doi: 10.1126/science.286.5445.1692

Frye, D., Zelazo, P. D., and Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cogn. Dev.* 10, 483–527. doi: 10.1016/0885-2014(95)90024-1

Garnham, W. A., and Ruffman, T. (2001). Doesn't see, doesn't know: is anticipatory looking really related to understanding of belief? *Dev. Sci.* 4, 94–100. doi: 10.1111/1467-7687.00153

Geiser, C. (2010). *Datenanalyse mit Mplus.* Wiesbaden: VS Verlag für Sozialwissenschaften. doi: 10.1007/978-3-531-92042-9

Happé, F. G. E. (1994). An advanced test of theory of mind:understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *J. Autism Dev. Disord.* 24, 129–154. doi: 10.1007/BF02172093

Henning, A., Hofer, T., and Aschersleben, G. (2012). "*Erweiterte "Theory of Mind-Skala" für 3- bis 11-jährige Kinder,*" in *Extended Theory-of-Mind Scale*, trans. Peterson, Wellman, and Slaughter. Saarbrücken: Universität des Saarlandes.

Hill, E. L. (2004). Executive dysfunction in autism. *Trends Cogn. Sci.* 8, 26–32. doi: 10.1016/j.tics.2003.11.003

Hollebrandse, B., van Hout, A., and Hendriks, P. (2012). First and second-order falsebelief reasoning: does language support reasoning about the belief of others? *Synthese* 191, 321–333. doi: 10.1007/s11229-012-0169-9

Hughes, C. (1998a). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Dev. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Hughes, C. (1998b). Finding your marbles: does preschoolers' strategic behavior predict later understanding of mind? *Dev. Psychol.* 34, 1326–1339. doi: 10.1037/0012-1649.34.6.1326

Hughes, C. (2001). "Executive dysfunction in autism: its nature and implications for the everyday problems experienced by individuals with autism," in *The Development of Autism*, eds J. Burack, T. Charman, N. Yirmiya, and P. Zelazo (Mahwah, NJ: Lawrence Erlbaum Associates), 255–274.

Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant Child Dev.* 20, 251–271. doi: 10.1002/icd.736

Hughes, C., and Ensor, R. (2007). Executive function and theory of mind: predictive relations from ages 2 to 4. *Dev. Psychol.* 43, 1447–1459. doi: 10.1037/0012-1649.43.6.1447

Hughes, C., and Russell, J. (1993). Autistic children's difficulty with mental disengagement from an object: its implications for theories of autism. *Dev. Psychol.* 29, 498–510. doi: 10.1037/0012-1649.29.3.498

Jahromi, L. B., and Stifter, C. A. (2008). Individual differences in preschoolers' self-regulation and theory of mind. *Merrill Palmer Q.* 54, 125–150. doi: 10.1353/mpq.2008.0007

Keenan, T., Olsen, D. R., and Marini, Z. (1998). Working memory and children's developing understanding of mind. *Aust. J. Psychol.* 50, 76–82. doi: 10.1080/00049539808257537

Kerr, A., and Zelazo, P. D. (2004). Development of "hot" executive function: the children's gambling task. *Brain Cogn.* 55, 148–157. doi: 10.1016/S0278-2626(03)00275-6

Kloo, D., and Perner, J. (2003). Training transfer between card sorting and false belief understanding: helping children apply conflicting descriptions. *Child Dev.* 74, 1823–1839. doi: 10.1046/j.1467-8624.2003.00640.x

Lalonde, C. E., and Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas Psychol.* 20, 163–198. doi: 10.1016/S0732-118X(02)00007-7

McAlister, A. R., and Peterson, C. C. (2013). Siblings, theory of mind, and executive functioning in children aged 3-6 years: new longitudinal evidence. *Child Dev.* 84, 1442–1458. doi: 10.1111/cdev.12043

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychol. Bull.* 135, 749–773. doi: 10.1037/a0016854

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moses, L. J. (2001). Executive accounts of theory of mind development. *Child Dev.* 72, 688–690. doi: 10.1111/1467-8624.00306

Müller, U., Liebermann-Finestone, D. P., Carpendale, J. I. M., Hammond, S. I., and Bibok, M. B. (2012). Knowing minds, controlling actions: the developmental relations between theory of mind and executive function from 2 to 4 years of age. *J. Exp. Child Psychol.* 111, 331–348. doi: 10.1016/j.jecp. 2011.08.014

Muthén, L. K., and Muthén, B. O. (1998–2012). *Mplus User's Guide*, 7th Edn. Los Angeles, CA: Muthén & Muthén.

Pacherie, E. (1997). "Motor-images, self-consciousness and autism," in *Autism as an Executive Disorder*, ed. J. Russell (Oxford: Oxford University Press), 215–255.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: clues to developmental primacy. *Dev. psychol.* 43, 974–990. doi: 10.1037/0012-1649.43.4.974

Perner, J. (1998). "The meta-intentional nature of executive functions and theory of mind," in *Language and Thought: Interdisciplinary Themes*, eds P. Carruthers and J. Boucher (Cambridge: Cambridge University Press), 270–283.

Perner, J., Kain, W., and Barchfeld, P. (2002). Executive control and higher-order theory of mind in children at risk of ADHD. *Infant Child Dev.* 11, 141–158. doi: 10.1002/icd.302

Perner, J., and Lang, B. (1999). Development of theory of mind and executive control. *Trends Cogn. Sci.* 3, 337–344. doi: 10.1016/S1364-6613(99)01362-5

Perner, J., and Lang, B. (2000). "Theory of mind and executive function: is there a developmental relationship," in *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*, eds S. Baron-Cohen, T. Tager-Flusberg, and D. Cohen (Oxford: Oxford University Press), 150–181.

Perner, J., and Wimmer, H. (1985). "John thinks that Mary thinks that..." attribution of second-order beliefs by 5- to 10-year-old children. *J. Exp. Child Psychol.* 39, 437–471. doi: 10.1016/0022-0965(85)90051-7

Petermann, F., and Petermann, U. (ed.). (2007). *Hamburg-Wechsler-Intelligenztest für Kinder IV (HAWIK-IV)*. Göttingen: Hogrefe.

Putko, A. (2009). Links between theory of mind and executive function: towards a more comprehensive model. *Pol. Psychol. Bull.* 40, 156–162. doi: 10.2478/s10059-009-0010-6

Razza, R. A., and Blair, C. (2009). Associations among false-belief understanding, executive function, and social competence: a longitudinal analysis. *J. Appl. Dev. Psychol.* 30, 332–343. doi: 10.1016/j.appdev.2008.12.020

Röthlisberger, M., Neuenschwander, R., Michel, E., and Roebers, C. M. (2010). Exekutive Funktionen: Zugrundeliegende kognitive Prozesse und deren Korrelate bei Kindern im späten Vorschulalter. *Z. Entwicklungspsychol. Pädagog. Psychol.* 42, 99–110. doi: 10.1026/0049-8637/a000010

Russell, J. (1996). *Agency: Its Role in Mental Development*. Hove: Lawrence Erlbaum Associates.

Russell, J. (1997). "How executive disorders can bring about an adequate theory of mind," in *Autism as an Executive Disorder*, ed. J. Russell (Oxford: Oxford University Press), 256–304.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., and Lee, K. (2006). The development of executive functioning and theory of mind: a comparison of chinese and U.S. preschoolers. *Psychol. Sci.* 17, 74–81. doi: 10.1111/j.1467-9280.2005.01667.x

Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol. Res. Online* 8, 23–74.

Schneider, W., Lockl, K., and Fernandez, O. (2005). "Interrelationships among theory of mind, executive control, language development, and working memory in young children: a longitudinal analysis," in *Young Children's Cognitive Development: Interrelationships Among Executive Functioning, Working Memory, Verbal Ability, and Theory of Mind*, eds W. Schneider, R. Schumann-Hengsteler, and B. Sodian (Mahwah, NJ: Lawrence Erlbaum Associates), 259–284.

Sebastian, C. L., Fontaine, N. M. G., Bird, G., Blakemore, S.-J., De Brito, S. A., McCrory, E. J. P., et al. (2012). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Soc. Cogn. Affect. Neurosci.* 7, 53–63. doi: 10.1093/scan/nsr023

Stevens, J. P. (2001). *Applied Multivariate Statistics for the Social Sciences*, 4th Edn. Mahwah, NJ: Lawrence Erlbaum Associates.

Tager-Flusberg, H., Sullivan, K., and Boshart, J. (1997). Executive functions and performance on false belief tasks. *Dev. Neuropsychol.* 13, 487–493. doi: 10.1080/87565649709540689

Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E. V., and Kliegel, M. (2013). Development of affective theory of mind across adolescence: disentangling the role of executive functions. *Dev. Neuropsychol.* 38, 114–125. doi: 10.1080/87565641.2012.733786

Völlm, B. A., Taylor, A. N. W., Richardson, P., Corcoran, R., Stirling, J., McKie, S., et al. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage* 29, 90–98. doi: 10.1016/j.neuroimage.2005.07.022

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and the constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128. doi: 10.1016/0010-0277(83)90004-5

Yang, J., Zhou, S., Yao, S., Su, L., and McWhinnie, C. (2009). The relationship between theory of mind and executive function in a sample of children from mainland china. *Child Psychiatry Hum. Dev.* 40, 169–182. doi: 10.1007/s10578-008-0119-4

Yu, C. Y. (2002). *Evaluating Cutoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes*. Los Angeles: University of California.

Zelazo, P. D., and Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: development and plasticity. *Child Dev. Perspect.* 6, 354–360. doi: 10.1111/j.1750-8606.2012.00246.x

Zimmermann, P., Gondan, M., and Fimm, B. (2002). *KITAP Testbatterie zur Aufmerksamkeitsprüfung für Kinder* [KITAP Attention Test Battery for Children]. Herzogenrath: Psychologische Testsysteme.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 February 2014; accepted: 08 June 2014; published online: 24 June 2014.
Citation: Austin G, Groppe K and Elsner B (2014) The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. Front. Psychol. ***5****:655. doi: 10.3389/fpsyg.2014.00655*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Austin, Groppe and Elsner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 30 May 2014
doi: 10.3389/fpsyg.2014.00523

Development of reference assignment in children: a direct comparison to the performance of cognitive shift

***Taro Murakami*[1]* and *Kazuhide Hashiya*[2]**

[1] *Department of Human-Environment Studies, Kyushu University, Fukuoka, Japan*
[2] *Faculty of Human-Environmental Studies, Kyushu University of Fukuoka, Japan*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Yui Miura, Kanazawa University, Japan
Gabriella Airenti, University of Torino, Italy

****Correspondence:***
Taro Murakami, Department of Human-Environment Studies, Kyushu University, 6-19-1 Hakozaki, Higashi-Ku, Fukuoka 812-8581, Japan
e-mail: taro.village@gmail.com

The referent of a deictic embedded in a particular utterance or sentence is often ambiguous. Reference assignment is a pragmatic process that enables the disambiguation of such a referent. Previous studies have demonstrated that receivers use social-pragmatic information during referent assignment; however, it is still unclear which aspects of cognitive development affect the development of referential processing in children. The present study directly assessed the relationship between performance on a reference assignment task (Murakami and Hashiya, in preparation) and the dimensional change card sort task (DCCS) in 3- and 5-years-old children. The results indicated that the 3-years-old children who passed DCCS showed performance above chance level in the event which required an explicit (cognitive) shift, while the performance of the children who failed DCCS remained in the range of chance level; however, such a tendency was not observed in the 5-years-old, possibly due to a ceiling effect. The results indicated that, though the development of skills that mediate cognitive shifting might adequately explain the explicit shift of attention in conversation, the pragmatic processes underlying the implicit shift, which requires reference assignment, might follow a different developmental course.

Keywords: preschooler, pragmatics, inference, reference assignment, executive function

INTRODUCTION

The referent of a deictic embedded in an utterance or sentence is often ambiguous. We communicate with others by interpreting the intended referent embedded in an utterance. However, interpreting another's referential intention is hardly achieved by a simple decoding process (Sperber and Wilson, 1986). The receiver must identify the intended referent based on a preceding situation or context. Reference assignment is a pragmatic process that enables disambiguation of a referent.

Previous studies have demonstrated that by age 2, children begin to use various non-verbal cues to determine the referent, such as the focus of the other person's attention (Baldwin, 1991), previous interactions with the other (Moll and Tomasello, 2007; Moll et al., 2007), the other's expression of preference (Repacholi, 1998), or the other's expression of glee or disappointment (Tomasello and Burton, 1994). Other researches have further demonstrated that children of the same age interpret an ambiguous request for absent objects, such as "Can you give *it* for me?" (Ganea and Saylor, 2007) or "Where's the ball?" (Saylor and Ganea, 2007), by reflecting on previous interactions with the experimenter that concerned particular objects. These studies agree in the sense that 2-years-old children have acquired the ability to use the relevant non-verbal information that has been gained through previous triad communications (self-object-other) in the process of interpreting an ambiguous referent.

Clark and Marshall (1981) pointed out the importance of linguistic evidence in processes where the receiver uses some form of information in interpreting a referent. Linguistic evidence could be termed as what the two persons have jointly heard, said, or are now jointly hearing as participants in the same conversation (also see Clark et al., 1983). In particular, the receiver must use contextual information from a shared conversational background to interpret the anaphoric expressions. With regard to the development of this ability, Ganea and Saylor (2007) demonstrated that 15- and 18-month-olds used the speaker's previous reference to an absent object to interpret the request.

However, in verbal communication, contextual redundancy often results in ambiguous referent interpretation because an object inevitably contains multiple aspects of information (name of object, color, function, and so on). When the labeling situation becomes ambiguous and the child has to determine from three or more alternatives which object is being labeled, 2-years-old interpret the novel words based on prior shared experiences with the experimenter (Akhtar et al., 1996; Diesendruck et al., 2004; Grasmann et al., 2009). Our previous study also indicated that 3-years-old children do not always use linguistic information from prior conversations retrospectively as a cue to interpret an ambiguous "How about this?" utterance (Murakami and Hashiya, in preparation). In this "reference assignment" task, 3-years-old children did not (though 5-years-old children did) refer retrospectively to the preceding linguistic context to identify the referent of an ambiguous utterance in the situation where the aspect to be referred in conversation was systematically changed (from shape to color or vice versa). The 3-years-old children, relative to 5-years-old, were also less proficient at shifting the referential aspect explicitly.

To effectively disambiguate an ambiguous referent, the receiver must attend to the same aspect as the sender. Evidence suggests that the ability to attend based on a verbal instruction might depend on the ability to perform a cognitive shift (directing attention from one aspect to another) (Murakami and Hashiya, in preparation). If the ability to interpret the ambiguous referent is based on the ability to track the interactions with the other, one could predict that children who are better at shifting their focus of attention should assign the referent more effectively when reflection on prior interactions with the other is useful. Primarily because of the close correlation between performance on "mind-reading" tasks, like False Belief, and the DCCS, the common underlying mechanism in terms of executive function (EF) is regarded as "domain-general" ability. To further examine this "domain-general" hypothesis, it should be determined whether EF predicts referent disambiguation performance. However, the relationship between these abilities has not yet been examined. Therefore, the present study directly assessed the association between reference assignment task and dimensional change card sort (DCCS) task performance in 3- and 5-years-old children.

The relationship between EF and mind-reading, as assessed in the False Belief task, has drawn many researchers' attention. In particular, DCCS performance, or cognitive shift, is significantly related to performance on the Contents False Belief task (Frye et al., 1995), even after controlling for individual differences in verbal ability (Carlson and Moses, 2001). It has been suggested that EF plays a central role in Theory of Mind development. In the False Belief task, the ability to perform a cognitive shift might be necessary to understand others' mental states based on a third-party situation. A related question is whether children better able to perform a cognitive shift would more effectively disambiguate the informative intention of a conversational partner.

The aims of the present study were to investigate the relationship between the ability to follow an explicit topic shift and the ability to perform a cognitive shift as measured by the DCCS. In addition, to appropriately assign the ambiguous referent, the receiver was required to follow the preceding context in accordance with the partner. We specifically examined whether children who were able to perform the cognitive shift necessary to follow another's attention would assign the appropriate referent to the ambiguous utterance. Therefore, we used reference assignment accuracy to investigate the development of disambiguation and cognitive shift ability.

MATERIALS AND METHODS

PARTICIPANTS

A total of 44 children (3-years-old: 11 girls, 9 boys; $M = 42.5$ months, $SD = 3.20$ months, and 5-years-old: 13 girls, 11 boys; $M = 66.2$ months, $SD = 3.71$ months) participated in this experiment. None of the children had participated in our previous study (Murakami and Hashiya, in preparation). All of the participants were born full-term and were healthy at the time of the study. Informed consent was obtained from the parents of all the children who participated. An additional four children who were 3 years of age were tested, but excluded from the final sample for the following reasons: understanding of color names was not confirmed (1), obvious bias when answering the questions (100% shape: 1, and 100% color: 1), and noncompliance with the reference assignment task (1).

MATERIALS AND DESIGN

Participants were tested individually in a room in the daycare center or preschool they attended. After establishing a rapport with the experimenter, the child participated in a test session. In a test session, the reference assignment task was always presented first. The entire experimental session lasted about 15 min, and all sessions were video recorded.

Reference assignment task

Stimuli. Laminated cards (14.8 × 21 cm) were used as stimuli. Each card represented one of five kinds of illustrations (umbrella, shoe, chair, cup, or car) painted in one of four colors (red, blue, yellow, or green). One stimulus set included all possible combinations of the objects and colors for a total of 20 cards (five shapes × four colors).

Procedure. One test session of the reference assignment task consisted of four trials. A trial consisted of five events, each of which included an explicit question (EQ) or an implicit question (IQ). In an EQ, participants were asked about either the shape or the color of the illustration on the card ["What's (the name of) this?" or "What color is this?"]. In an IQ, participants were asked, "How about this?" The sequence of events included in a trial was as follows: the first event was always an EQ followed by an IQ (PreS-IQ). Another EQ (ESQ) was then asked, but the dimension (shape/color) differed. The ESQ was then followed by two IQs (PostS-IQ1, 2). Half of the four trials began with an EQ about the shape, whereas the other half of the trials began with an EQ about the color. The order of the trials was counterbalanced across participants.

The child was shown a card, and the experimenter said, "Now, let's try a game. Listen to me carefully and answer the questions." The experimenter continued to ask questions one at a time about the five cards (see **Figure 1**). The experimenter made eye contact with the children, and nodded regardless of whether the child had correctly answered the question(s). After asking questions about the five cards, the experimenter aligned the cards in front of the child to indicate to the child that one trial had been completed. The experimenter then took out a new set of cards and began the next trial. A total of four trials were conducted with each child.

Scoring. Responses for each trial were coded on a dichotomous rating, defined as follows. For EQs, an appropriate answer was coded as 1, and an incorrect answer was coded as 0 (e.g., an answer that referred to the "color" aspect when the child was asked about an object's "shape" was scored as 0). For IQs, the retrospective answer that referred to the dimension of the explicit question asked just before the implicit question was coded as 1. In addition, a coding battery was applied to the analysis in order to describe the sequential pattern of the child's response beyond a single event.

Base-Assignment Score. When both the EQ and PreS-IQ were coded as 1, the Base-Assignment score was coded as 1, reflecting

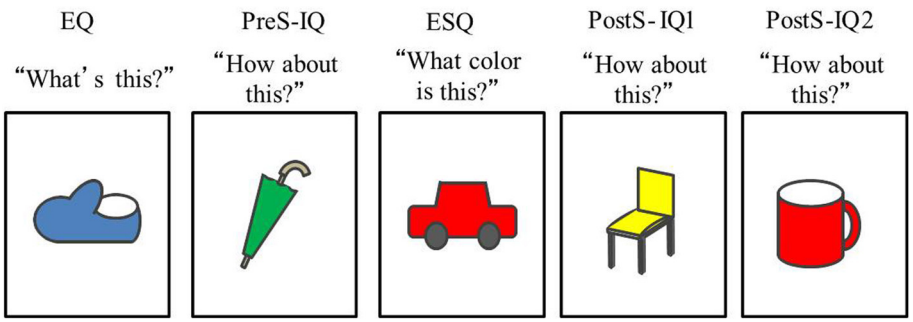


FIGURE 1 | Schematic sequence of an event in the Reference Assignment task, which includes 5 events in a fixed order of EQ/PreS-IQ/ESQ/PostS-IQ1/PostS-IQ2.

that the child had "appropriately" identified the reference in the absence of a topic shift.

Shift Score. The Shift score indicates a child's ability to switch to the explicit question; therefore, this score was coded as 1 when both the EQ and ESQ were coded as 1.

Re-Assignment Score. The Re-Assignment score denotes a child's referential assignment based on topic shift; therefore, the score was coded as 1 when both the ESQ and Post-IQ1 were coded as 1.

Follow-Re-Assignment (Follow-RA) Score. The Follow-RA score indicates whether the child interpreted the repetition of the same ambiguous question consistently; therefore, the score was coded as 1 when both the PostS-IQ1 and PostS-IQ2 were coded as 1.

Dimensional change card sort task

The procedure of the DCCS was consistent with that of Kirkham et al. (2003).

Stimuli. The model cards consisted of two white laminated cards (10.5 × 7.5 cm); one card depicted a red truck and the other depicted a blue star. The sorting cards were the same size and shape as the model cards, but each depicted a blue truck or red star. Thus, no sorting card matched a model card in both color and shape. A sorting card was mounted over the bin of each box. The children were trained to sort by color with training cards that depicted blue or red caps, and were trained to sort by shape with training cards that depicted yellow cars or stars.

Procedure. The child was shown the two sorting boxes with the model cards. The experimenter then introduced the child to the training part of the game, which consisted of sorting cards that were similar in only one dimension (i.e., cards depicting blue and red caps for the color game or cards depicting yellow cars and stars for the shape game).

The first dimension on which children were trained was counterbalanced across children within each age × gender. Each child was given between 4 and 8 cards (i.e., allowing for four errors). Two cards of one dimension were presented first, followed by two cards of another dimension. Children had to correctly sort four cards (two for each dimension) to pass the training phase. Feedback was provided to the children. The last dimension sorted during the training phase was always the first dimension administered during the test trials (e.g., if the final training card sorted depicted red caps, then the first test dimension would be color). The test trials started immediately after the child had completed the training trials.

There was a minimum of 12 test trials (i.e., six consecutive trials for the first dimension, and six consecutive trials for the second dimension). Because children were required to sort six trials in a row to reach criterion, additional trials were administered until the child passed criterion for that dimension. Additional trials were needed on only two occasions: two 3-years-old children required 6 (1) and 7 trials (1), and one 5-years-old child required 8 trials to reach criterion on the first dimension. The same pseudo-random order of card presentation was used for all children. Before each trial, the child was asked to tell the experimenter the rules of the current game by pointing to the appropriate boxes in response to "knowledge" questions (e.g., "Where do the red ones go in the color game? Where do the blue ones go?"). During alternating trials, the experimenter typically stated the rules and had the child answer the knowledge questions. We randomly varied the value (e.g., red or blue) that was mentioned first.

Children were given feedback on their response to the knowledge question. If the child's response was correct, the experimenter said, "Excellent!" or "Very good." The child was then given the next card and asked to sort it according to the appropriate dimension (e.g., "Here's a blue one. Where does it go?" or "Here's a car. Where does it go?"). If the child answered the knowledge question incorrectly, the experimenter restated the rules and asked the knowledge question again. If the child responded incorrectly again, the error was noted and the next trial commenced.

Note that the experimenter indicated only the relevant dimension of each stimulus ("Here's a blue one"), whereas in their early work, Zelazo et al. (1996) labeled both dimensions of each stimulus ("Here is a blue car"). In addition, feedback was not provided to the child during testing. The child was asked to place the sorting cards face down in the sorting boxes. After the child had correctly sorted six cards by the first dimension, the sorting dimension was switched. Moreover, children were allowed to self-correct.

Then, based on their DCCS performance, children were divided into two groups: DCCS-passed and DCCS-failed. To pass the DCCS, children must correctly sort five of the six cards. We examined whether the children who passed the DCCS showed better performance on the reference assignment task than the children who failed the DCCS; therefore, we used this classification as a categorical factor on the reference assignment task.

RESULTS

REFERENCE ASSIGNMENT TASK

For the reference assignment task, preliminary analysis revealed no gender differences or effect of trial order; thus, these factors were collapsed in the subsequent analyses. **Table 1** shows the mean score for each event in the reference assignment task. The averaged score for each event was compared in a 2×2 ANOVA with Age (3 vs. 5 years) as a between-subjects factor and Event (Base-Assignment vs. Shift vs. Re-Assignment vs. Follow-RA) as a within-subjects factor. The results revealed a significant interaction between Age and Event [$F_{(3,\ 126)} = 3.71$, $p = 0.01$, $\eta_p^2 = 0.08$] and a significant main effect of Age and Event [Age: $F_{(1,\ 42)} = 22.93$, $p < 0.001$, $\eta_p^2 = 0.35$; Event: $F_{(3,\ 126)} = 18.07$, $p < 0.001$, $\eta_p^2 = 0.30$]. Multiple comparisons revealed that 5-year-old outperformed 3-years-old except in the Base-Assignment score ($p < 0.01$ at maximum). Moreover, in 3-years-old, Base-Assignment ($M = 3.2$, $SD = 0.88$) included more "appropriate" answers than other scores on Shift ($M = 2.4$, $SD = 1.19$), Re-Assignment ($M = 2.0$, $SD = 0.65$) and Follow-RA ($M = 2.0$, $SD = 0.83$), $p < 0.01$ at maximum. In 5-years-old, Base-Assignment ($M = 3.4$, $SD = 0.78$) and Shift ($M = 3.7$, $SD = 0.55$) included more "appropriate" answers than Re-Assignment ($M = 2.8$, $SD = 0.87$) and Follow-RA ($M = 2.7$, $SD = 0.86$), $p < 0.01$ at maximum. The age-dependent patterns observed in the present study are consistent with those of our previous study (Murakami and Hashiya, in preparation).

To examine whether there was a bias to respond to a specific aspect when presented with ambiguous questions, we tallied the number of errors for shifts from color to shape (0–2), or shape to color, during each event for the two age groups. **Figure 2** shows the mean error for IQs in 3- and 5-years-old. The results of a t-test for each event suggested there was no difference in 3-years-old children [PreS-IQ, $t_{(17)} = 1.000$, $p = 0.33$, $r = 0.24$; PostS-IQ1, $t_{(17)} = 0.579$, $p = 0.57$, $r = 0.14$; PostS-IQ2, $t_{(17)} = 0.437$, $p = 0.66$, $r = 0.11$]; however, 5-years-old children were more likely to answer in the shape than the color when an ambiguous question was presented [PreS-IQ, $t_{(23)} = 2.632$, $p = 0.01$, $r = 0.48$; PostS-IQ1, $t_{(23)} = 2.077$, $p = 0.049$, $r = 0.40$; PostS-IQ2, $t_{(23)} = 1.967$, $p = 0.06$, $r = 0.38$]. The response bias observed in 5-years-old is inconsistent with our previous research (Murakami and Hashiya, in preparation). Although 5-years-old tended to state the shape of the object in response to an ambiguous question, the exact error rate (29–40%) remained within the range of chance; thus, the results may not have necessarily indicated a shape bias (Landau et al., 1988). Therefore, we did not consider this a significant reaction characteristic of 5-years-old and continued the analysis.

DIMENSIONAL CHANGE CARD SORT TASK

For the DCCS, all children sorted all the cards correctly for the first dimension. After the switch to the second dimension, 93% of the children consistently sorted all the cards correctly, or all incorrectly. Given the lack of variance, nonparametric categorical analyses (chi-square) were used to analyze the data. The number of children who successfully switched dimensions in the card-sorting task is shown in **Table 2**. The majority of 3-years-old performed poorly (only 25% successfully switched dimensions), while most 5-years-old performed well (66% successfully switched dimensions). The difference in performance between

Table 1 | Mean score and standard deviation (in parentheses) for each event of the Referential Assignment task.

	Base-assignment	Shift	Re-assignment	Follow-RA
3-years-old	3.2 (0.88)	2.4 (1.19)	2.0 (0.65)	2.0 (0.83)
5-years-old	3.4 (0.78)	3.7 (0.55)	2.8 (0.87)	2.7 (0.86)

Table 2 | Distribution of the group of age × performance on DCCS.

Age	DCCS	*N* (girls)	Mean months	*SD*
3	Failed	15 (8)	42.2	3.3
	Passed	5 (3)	43.2	2.9
5	Failed	8 (6)	64.9	4.4
	Passed	16 (7)	66.9	3.3

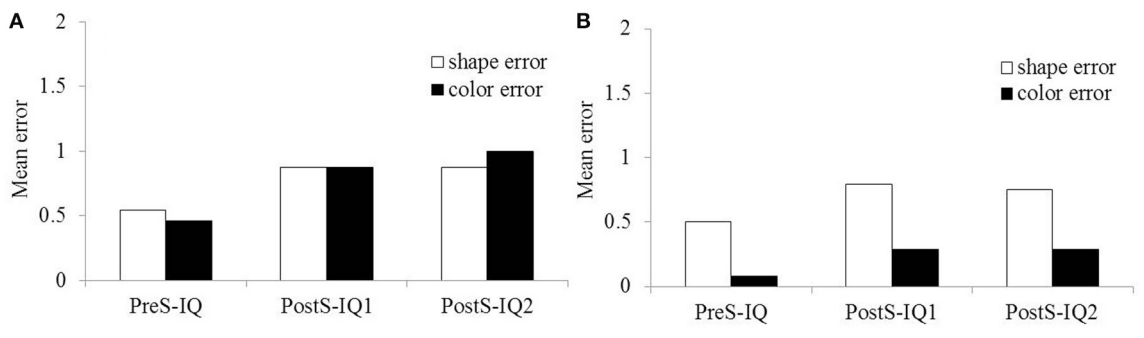


FIGURE 2 | Mean error for Implicit Questions in (A) 3-years-old and (B) 5-years-old.

the two groups was significant [$\chi^2_{(df=1,\ N=44)} = 6.013, p < 0.05$]. These results suggest that the current sample was similar to those of previous studies.

COMPARISON BETWEEN THE REFERENCE ASSIGNMENT TASK AND DCCS TASK

The number of "appropriate" responses in the reference assignment task was analyzed using a 2 × 2 × 4 mixed ANOVA with Age (3 vs. 5 years) and DCCS group (passed vs. failed) as between-subjects factors, and Event (Base-Assignment vs. Shift vs. Re-Assignment vs. Follow-RA) as a within-subjects factor. No significant interactions between factors were found (see **Figure 3**); however, main effects of Age and Event [Age: $F_{(1,\ 40)} = 16.48, p < 0.001, \eta_p^2 = 0.28$; Event: $F_{(3,\ 120)} = 16.59, p < 0.001, \eta_p^2 = 0.34$] were observed. The main effect of DCCS was not significant.

To determine the rate of correct responses to the questions, the proportion of appropriate responses was compared with chance levels (=2). For the 3y-failed group, one-sample *t*-tests indicated that performance was above chance level for the Base-Assignment score [$t_{(14)} = 6.00,\ p < 0.001,\ r = 0.85$], but performance in other events remained within the range of chance. One-sample *t*-tests for the 3y-passed group indicated that performance was above chance level only for the Shift questions [$t_{(4)} = 3.16$, $p = 0.034,\ r = 0.85$]. On the other hand, analysis of 5y-failed group indicated that performance was above chance level for all events [Base-Assignment; $t_{(7)} = 4.25, p = 0.004, r = 0.85$; Shift; $t_{(7)} = 15.00, p < 0.0001, r = 0.99$; Re-Assignment; $t_{(7)} = 2.05$, $p = 0.08, r = 0.61$; Follow-RA; $t_{(7)} = 2.05, p = 0.095, r = 0.61$]. Analysis of the 5y-passed group also indicated that performance was above chance level for all events [Base-Assignment; $t_{(15)} = 7.90, p < 0.001, r = 0.90$; Shift; $t_{(15)} = 10.50, p < 0.001$, $r = 0.94$; Re-Assignment; $t_{(15)} = 4.34,\ p = 0.001,\ r = 0.75$; Follow-RA; $t_{(15)} = 3.50, p = 0.003, r = 0.67$].

DISCUSSION

The current study directly compared performance on a reference assignment task with DCCS performance in preschoolers, and identified a relationship between the ability to follow an explicit utterance and the ability to perform a cognitive shift, which develops between 3 and 5 years of age (Zelazo et al., 1996, 2003; Carlson and Moses, 2001; Kirkham et al., 2003; Müller et al., 2006; Moriguchi et al., 2007; Moriguchi and Hiraki, 2009). However, the present findings indicate that some aspects of the ability to disambiguate based on prior verbal exchanges do not always reflect a cognitive shift. A previous study showed that children interpret the ambiguous speech of others by referring to information from a prior situation in which one potential referent was salient (Murakami and Hashiya, in preparation). In the reference assignment task, children in the current study replicated this finding. Performance on the DCCS was also consistent with the previously observed patterns for these age groups. These results suggest that the participant group in the current study did not differ qualitatively from those of previous studies.

The comparison of these two tasks contributes to our knowledge of the relationship between EF and understanding verbal instruction. On the Shift score, although the ANOVA results did not show an Age × DCCS interaction, a comparison with chance level showed that the 3-years-old children who passed the DCCS effectively redirected their attention in response to explicit verbal instruction. These results suggest that the ability to focus on another aspect of a target in response to language is necessary to shift the classification rule, such as in the DCCS. However, even though they could shift their explicit attention, the 3-years-old children who passed the DCCS did not retrospectively assign the referent based on the preceding explicit verbal exchange. These results suggest that the cognitive ability

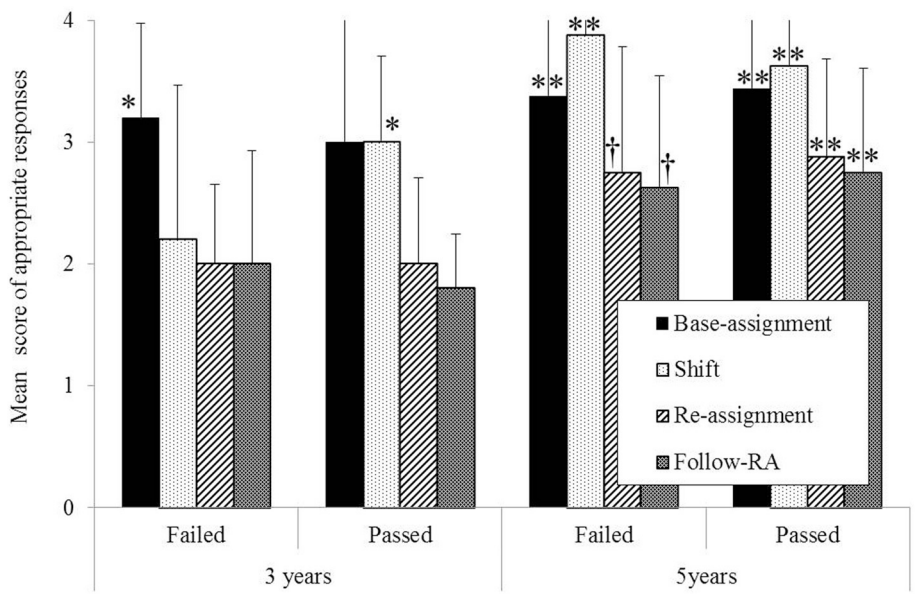


FIGURE 3 | Mean score of appropriate responses. **,* and †indicate that the score was above chance level (=2), $p < 0.01$, 0.05, and 0.10, respectively.

of shifting attention does not always facilitate the retrospective reference.

In a similar fashion, both groups of 5-years-old children showed only moderate performance in ESQ, even though it was above chance level. However, their verbal shifting performance seemed to show a ceiling effect. This inconsistency suggests that the difficulties in nonverbal shifting are not tightly related to verbal shifting ability, which might be consistent with previous findings about the knowledge questions of the DCCS (Kirkham et al., 2003), which are structurally similar to the ESQ in the reference assignment task. In addition, the ceiling effect in 5-years-old might be explained by a developmental improvement in sensitivity toward verbal instruction.

Moreover, the DCCS-failed group of 5-years-old children also showed a marginal tendency to retrospectively reference. When we compared Re-Assignment and Follow-RA scores with chance level, we found that both groups of 5-years-old children disambiguated the ambiguous deictic; they tended to interpret the ambiguous utterance retrospectively. These results suggest that even the children who failed the DCCS could disambiguate the ambiguous utterance.

The reference assignment task, which enables systemic assessment of one's understanding of a deictic, potentially represents a means of separating underlying systems that mediate the process of disambiguation. Further, the current results demonstrate that the ability of cognitive shift is correlated with the ability to disambiguate the linguistic referent, but only to a limited extent.

Thus, the results did not support the expectation that the ability of cognitive shift would entirely explain the ability to disambiguate a linguistic referent, but rather suggested independent development of retrospective referencing and cognitive shift.

The ability to use contextual information from a shared conversational background is one of the essential pragmatic skills (Clark and Marshall, 1981) in effectively inferring the references of another (Sperber and Wilson, 2002). Though previous findings have demonstrated that even 2-years-old infants interpret an ambiguous request for an object in terms of prior interactions with the requestor (Ganea and Saylor, 2007; Saylor and Ganea, 2007), the current study suggested the difficulty for 3-years-old children in identifying an ambiguous referent based only on verbal information. Considering these concerns, our results may imply that several extra processes are required for completing our reference assignment task: the processes such as acquisition of a semantic definition of the deictic, or the conventional principle that the ambiguous "this" embedded in a specific form of the sentence refers to some salient aspect or event expressed in the precedent utterance, should be the candidates for such missing pieces. Based on the current findings, the detailed interactions of such contributing factors should be a focus of future studies. Thus, studies geared toward dissecting the development of pragmatic communication might serve as an effective means of describing the generality and specificity of the development of EF, especially when the reference assignment task is included in the test battery.

ACKNOWLEDGMENTS

We sincerely thank H. Ohgami, N. Matsushima, and K. Ishikawa for their help in conducting the study. We also thank the preschools, parents, and children who participated in the research. This work was supported by Grant-in-Aid for Scientific Research on Innovative Areas # 25118003, "The Evolutionary Origin and Neural Basis of the Empathetic Systems" and (C) # 23500330.

REFERENCES

Akhtar, N., Carpenter, M., and Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Dev.* 67, 635–645. doi: 10.2307/1131837

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Dev.* 62, 874–890. doi: 10.1111/j.1467-8624.1991.tb01577.x

Carlson, S. M., and Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333

Clark, H. H., and Marshall, C. R. (1981). "Definite reference and mutual knowledge," in *Elements of Discourse Understanding*, eds A. K. Joshi, B. L. Webber, and I. A. Sag (New York, NY: Cambridge University Press), 10–63.

Clark, H. H., Schreuder, R., and Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *J. Verb. Learn. Verb. Behav.* 22, 245–258. doi: 10.1016/S0022-5371(83)90189-5

Diesendruck, G., Markson, L., Akhtarm N., and Reudor, A. (2004). Two-year-olds' sensitivity to speakers' intent: an alternative account of Samuelson and Smith. *Dev. Sci.* 7, 33–41. doi: 10.1111/j.1467-7687.2004.00320.x

Frye, D., Zelazo, P. D., and Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cogn. Dev.* 10, 483–527. doi: 10.1016/0885-2014(95)90024-1

Ganea, P. A., and Saylor, M. M. (2007). Infants' use of shared linguistic information to clarify ambiguous requests. *Child Dev.* 78, 493–502. doi: 10.1111/j.1467-8624.2007.01011.x

Grasmann, S., Stracke, M., and Tomasello, M. (2009). Two-year-olds exclude novel objects as potential referents of novel words based on pragmatics. *Cognition* 112, 488–493. doi: 10.1016/j.cognition.2009.06.010

Kirkham, N. Z., Cruess, L., and Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Dev. Sci.* 6, 449–476. doi: 10.1111/1467-7687.00300

Landau, B., Smith, L. B., and Jones, S. (1988). The importance of shape in early lexical learning. *Cog. Dev.* 3, 299–321. doi: 10.1016/0885-2014(88)90014-7

Moll, H., Carpenter, M., and Tomasello, M. (2007). Fourteen-month-olds know what others experience only in joint engagement. *Dev. Sci.* 10, 826–835. doi: 10.1111/j.1467-7687.2007.00615.x

Moll, H., and Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Dev. Psychol.* 43, 309–317. doi: 10.1037/0012-1649.43.2.309

Moriguchi, Y., and Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6017–6021. doi: 10.1073/pnas.0809747106

Moriguchi, Y. Lee, K., and Itakura, S. (2007). Social transmission of disinhibition in young children. *Dev. Sci*. 10, 481–491. doi: 10.1111/j.1467-7687.2007.00601.x

Müller, U., Dick, A. S., Gela, K., and Overton, W. F. (2006). Negative priming in preschoolers' flexible rule use on the dimensional change card sort task. *Child Dev.* 77, 395–412. doi: 10.1111/j.1467-8624.2006.00878.x

Repacholi, B. M. (1998). Infants' use of attentional cues to identify the referent of another person's emotional expression. *Dev. Psychol.* 34, 1017–1025. doi: 10.1037/0012-1649.34.5.1017

Saylor, M. M., and Ganea, P. (2007). Infants interpret ambiguous requests for absent objects. *Dev. Psychol.* 43, 696–704. doi: 10.1037/0012-1649.43.3.696

Sperber, D., and Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sperber, D., and Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind Lang.* 17, 3–23. doi: 10.1111/1468-0017.00186

Tomasello, M., and Burton, M. E. (1994). Learning words in non-ostensive contexts. *Dev. Psychol.* 30, 639–650. doi: 10.1037/0012-1649.30.5.639

Zelazo, P. D., Frye, D., and Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cog. Dev*. 11, 37–63. doi: 10.1016/S0885-2014(96)90027-1

Zelazo, P. D., Müller, U., Frye, D., and Marchovitch, S. (2003). The development of executive function in early childhood. *Monogr. Soc. Res. Child Dev.* 68, 131–151. doi: 10.1111/j.0037-976X.2003.00261.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 January 2014; accepted: 12 May 2014; published online: 30 May 2014.
Citation: Murakami T and Hashiya K (2014) Development of reference assignment in children: a direct comparison to the performance of cognitive shift. Front. Psychol. ***5****:523. doi: 10.3389/fpsyg.2014.00523*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Murakami and Hashiya. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 20 March 2014
doi: 10.3389/fpsyg.2014.00240

The role of executive functions in the pragmatic skills of children age 4–5

Bénédicte Blain-Brière[1], Caroline Bouchard[2]* and Nathalie Bigras[3]

[1] *Qualité Éducative des Services de Garde et Petite Enfance, Département de Psychologie, Université du Québec à Montréal, Montréal, QC, Canada*
[2] *Qualité Éducative des Services de Garde et Petite Enfance, Département D'études sur L'enseignement et L'apprentissage, Université Laval, Ville de Québec, QC, Canada*
[3] *Qualité Éducative des Services de Garde et Petite Enfance, Département de Didactique, Université du Québec à Montréal, Montréal, QC, Canada*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Taro Murakami, Kyushu-University, Japan
Mako Okanda, Kobe University, Japan

****Correspondence:***
Caroline Bouchard, Qualité Éducative des Services de Garde et Petite Enfance, Département D'études sur L'enseignement et L'apprentissage, Université Laval, 2320 rue des Bibliothèques (bureau 1234), G1V 0A6, Ville de Québec, QC, Canada
e-mail: caroline.bouchard@fse.ulaval.ca

Several studies suggest that pragmatic skills (PS) (i.e., social communication) deficits may be linked to executive dysfunction (i.e., cognitive processes required for the regulation of new and complex behaviors) in patients with frontal brain injuries. If impairment of executive functions (EF) causes PS deficits in otherwise healthy adults, could this mean that EF are necessary for the normal functioning of PS, even more so than cognitive maturation? If so, children with highly developed EF should exhibit higher levels of PS. This study aimed to examine the link between EF and PS among normally developing children. A secondary goal was to compare this relationship to that between intellectual quotient (IQ) and PS in order to determine which predictor explained the most variance. Participants were 70 French-speaking preschool children (3;10–5;7 years old). The PS coding system, an observational tool developed for this study, was used to codify the children's PS during a semi-structured conversation with a research assistant. Five types of EF processes were evaluated: self-control, inhibition, flexibility, working memory and planning. IQ was estimated by tallying the scores on a receptive vocabulary test and a visuoconstructive abilities test. The results of the test of differences between correlation coefficients suggest that EF contributed significantly more than IQ to the PS exhibited by preschoolers during conversation. More specifically, higher inhibition skills were correlated with a decrease in talkativeness and assertiveness. EF also appeared to foster quality of speech by promoting the ability to produce fluid utterances, free of unnecessary repetition or hesitation. Moreover, children with a high working memory capacity were more likely to formulate contingent answers and produce utterances that could be clearly understood by the interlocutor. Overall, these findings help us better understand how EF may assist children in everyday social interactions.

Keywords: pragmatic skills, communication, executive functions, vocabulary, visuoconstructive abilities, cognitive development, language acquisition, early childhood

INTRODUCTION

Pragmatic skills (PS) in children refer to the ability to use communication strategies in social interactions (Owens, 2011). These skills contribute to children's psychosocial adjustment and academic achievement (Ervin-Tripp, 1978; McKown, 2007; Coplan and Weeks, 2009; Brinkman et al., 2013). Russell and Grizzle (2008) examined 24 instruments used to assess PS among children and adolescents in order to identify the core domains of PS. They found just over 1000 different items in these instruments, which they grouped into 17 domains and further classified into three sets: (1) Precursors/enablers (e.g., nonverbal communication; discourse attentiveness and empathy; speech characteristics and fluency), (2) Basic exchanges/rounds (e.g., conversational turn taking; topic control and maintenance; requests), (3) Extended literal and non-literal discourse (e.g., negotiations, directions, and instructions; theory of mind; narrative; Gricean principles) (Russell and Grizzle, 2008). Although this classification is helpful, there is still no empirical finding corroborating such a categorization. In fact, Russell and Grizzle (2008) reported that almost none of the authors who constructed the instruments they inventoried had performed factorial analyses. Thus, in order to describe the empirical dimension of PS, specifically among preschoolers, the authors of the present study carried out a systematic literature review and performed a factor analysis (Blain-Brière et al., submitted). They concluded that preschoolers' PS can be divided into five categories: conversational complexity, talkativeness, assertiveness, communicative control and responsiveness.

Studying the development of this five categories of PS in preschooler, a year or two prior to the school commencement, is crucial because it is around this age that children start to play interactively with each other (Smith, 2003). Their ability to manifest PS will shape their early socialization experiences, influence their social acceptance and help them develop their socials skills (Black and Hazen, 1990; McKown, 2007). By preschool age, children have already mastered a wide range of PS (Adams, 2002).

By age 1, they know how to request something by pointing to it (Carpenter et al., 1998; Liszkowski et al., 2004). Between the age of 2 and 4, Martinez (1987) shows that children's speech contains more turnabout, namely a utterance that have the dual function of responding to the speaker and restarting the conversation. Pellegrini et al. (1987) note also that children of this age tend to exchange more utterances with their interlocutor, from around 14 utterances per minute at age 2 to about 22 utterances per minute at age 3–4. By about age 3, they can already adapt their speech to an interlocutor (Dunn and Kendrick, 1982). Sachs et al. (1991) showed that at age 3, children have a tendency to ask adults questions regardless if it is an appropriate time to do it, whereas most children by age 5 are able to wait until the adult has finished speaking before querying them. Some abilities, such as understanding figurative speech, are not completely acquired until adolescence or even adulthood (Nippold, 1985; Ervin-Tripp et al., 1990; Spector, 1996).

Developmental studies have thus shown that children are constantly required to manifest PS, and that these skills become increasingly cognitively demanding as they get older. Could cognitive factors therefore play a role in the acquisition of PS? For instance, before children are able to wait their turn to speak, surely they must first acquire the ability to inhibit a response. In order words, inhibition skills, a cognitive process involved in executive functioning, would need to be sufficiently developed before a child could refrain from speaking during his interlocutor's speaking turn. In brief terms, executive functions (EF) are defined as the mechanisms that regulate cognition by modulating the operation of a variety of cognitive processes including inhibition, but also working memory (WM), flexibility and planning (Lehto et al., 2003; Blair et al., 2005). Yet, while the involvement of cognitive processes such as EF in PS seems logical, to date, few authors have investigated this relationship among typically developing children.

In adults, PS deficits (e.g., excessive talkativeness, subject shifting, problems understanding indirect questions) following a prefrontal brain injury are well-documented in the literature (Martin and McDonald, 2003; Douglas, 2010; Dardier et al., 2011). Several studies have found that PS deficits are correlated with executive dysfunction in patients with traumatic brain injury (TBI) (McDonald and Pearce, 1996, 1998; Channon and Watts, 2003; Douglas, 2010). This correlation implies that EF are necessary for the normal functioning of PS. Based on this premise, it seems probable that EF may also contribute to the acquisition of PS in normally developing children (Blain-Brière et al., submitted). Therefore, children with well-developed EF should exhibit better PS than other peers of the same age. Of course, these deductions are theoretical and need to be proven. Yet, there is evidence supporting them. For instance, children with executive dysfunction, caused by a neurodevelopmental disease such as attention deficit hyperactivity disorder (ADHD) (Humphries et al., 1994; Bruce et al., 2006) or autism (Ozonoff, 2001; Norbury et al., 2004; Bishop and Norbury, 2005; Reisinger, 2011; Schuh, 2012), have been found to exhibit PS deficits.

Even among normally developing children, according to Nilsen and Graham (2009) and Schuh (2012), there is proof of a correlation between EF and PS. To evaluate PS, these authors used a similar referential communication experimental protocol that specifically measured how children used speech to signify things in the world. In this task, the participant was typically asked by the examiner to choose an object from an array of objects. The participant had to take into account the context of the situation such as what the examiner could see from his position. For instance, if the examiner could not see the red object from where he was standing, the participant would conclude that the object asked for was not red. In their study among typically developing children aged 3–5 years, Nilsen and Graham (2009) noted that inhibition contributed to the children's ability to consider the perspective of the examiner when choosing the right object. Their interpretation was that inhibition allowed the children to inhibit their own perspective in order to consider the viewpoint of the examiner (Nilsen and Graham, 2009). Schuh (2012) also used a referential communication task to study the influence of WM among typically developing children aged 8–17 years. She demonstrated that children with a higher WM capacity responded more accurately to their partner's request because they were able to take into account information that the latter did not know about the situation. The results of Nilsen and Graham (2009) and Schuh (2012) show that inhibition and WM may increase the ability to interpret the perspective of others. Consequently, children with highly developed EF may be better at grasping the speech of their interlocutor, especially when it is ambiguous, and respond accordingly. This gain in responsiveness during conversation could mean that EF increase PS among children. However, as pointed out by Bishop and Adams (1991), referential communication tasks are not necessarily representative of how children communicate in an unstructured conversational setting. These authors demonstrated, for instance that children who provided excessive and irrelevant information in such a task did not act the same way during open-ended conversation. Hence, the link between PS and EF needs to be demonstrated in a more natural context in order to confirm that EF truly benefit children in conversation. To date, very few studies have examined the relationship between EF and PS through a direct observation measure of PS (Jagot et al., 2001). An observational research design is needed to confirm that children do indeed rely on EF in their everyday social interactions.

Moreover, it is important to note that EF are not the only cognitive processes thought to contribute to PS. In fact, previous research has shown that vocabulary, visuoconstructive abilities and intellectual quotient (IQ) may also be related to PS (McDonald, 2000; Bonifacio et al., 2007; McKown, 2007). Nevertheless, regression analyses have demonstrated that EF may make a unique contribution to the PS of children, even after controlling for vocabulary size and age (Nilsen and Graham, 2009). However, while regression analyses may prove that EF explain a unique part of the variance, they cannot tell which predictor, among vocabulary, visuoconstructive abilities, IQ and EF, has the strongest relationship with PS. On the other hand, a test of differences between correlations would make it possible to determine the relative role played by each predictor and whether these differences are significant. Such an analysis would allow answering the question of (1) whether overall cognitive maturation (e.g., vocabulary, visuoconstructive abilities and IQ) has more or less

the same influence on PS as EF or (2) whether each EF process plays a specific role in PS which is significantly different from that played by other cognitive processes.

The above-cited TBI and general population studies have another shortcoming when it comes to demonstrating a link between PS and EF. They usually use a very limited number of measures of PS and/or EF. Douglas (2010), for instance, measured only EF in the verbal domain [verbal fluency (FAS), verbal memory (RAVLT) and speed and capacity of language-processing (SCOLP)]. Nilsen and Graham (2009) and Schuh (2012), for their part, measured only PS related to referential communication. Consequently, these authors could not show exactly how each EF process may contribute to each PS separately.

To further our understanding of the possible role played by EF in the normal acquisition of PS, this study aimed to examine the link between EF and PS among typically developing preschoolers. This study was innovative insofar as it used a direct observational tool to evaluate PS in order to assess how EF might influence the PS of children in their everyday social interactions. Moreover, a test of differences between correlations helped us to understand to what extent the link between EF and PS is different from the relationship between an IQ estimate and PS. This study also adds to previous work in the field by using a wide range of variables to measure PS (14 variables) and EF (self-control, inhibition, WM, flexibility and planning).

MATERIALS AND METHODS

PARTICIPANTS

The study sample consisted of 70 French speaking children (34 girls and 36 boys) with an average age of 4 and a half years (55.2 months, $SD = 4.5$ months, 3;10–5;7 years). They were all recruited from a subsidized childcare center in a class designed for children who will enter the school system in a year or two. In order to participate, the children's language had to be developing normally based on the information reported by their childcare provider and the results of a receptive vocabulary task. Eighty children were initially recruited, 10 of whom could not be included in the study, either because of suspected language delays (4 subjects), because the child was absent when the testing took place (3 subjects) or as a result of technical problems during the video recording of the conversation sample (3 subjects). As for the sociodemographic characteristics of the participants, 30.6% lived in a household with an income of less than $30,000, while the household income for 28.7% was $30,000–70,000, for 28.1% was $70,000–100,000, and for 28.1% exceeded the threshold of $100 000. As for the level of education of the participants' mothers, 3.1% of mothers had not completed high school (11 grades in Quebec, Canada), 9.4% had at most a high school education, 12.5% had a vocational school diploma, 26.6% had a college education and 48.4% had a university degree.

MATERIALS

Pragmatic skills

The *Grille d'observation des habiletés pragmatiques des enfants d'âge préscolaire* (Pragmatic Skills Coding System—Preschool Version (PSCS-P) (Blain-Brière et al., submitted) was used to measure PS. This instrument was developed after three years of research by the authors of this article in order to palliate for the lack of validated observational tools for assessing PS among preschoolers. The PSCS-P measures 14 PS parameters during a semi-structured conversation with an examiner. The parameters were developed by selecting variables from 21 utterance coding systems, themselves retrieved from a systematic literature review. To ensure content quality of the parameters selected, independents expert's advices were solicited and factor analysis were performed. **Table 1** describes how these variables were codified and their Intraclass correlation coefficient (ICC) measured in the validation process on a sample of 18 participants. It also presents the five scales that they are associated with and their coherence coefficients. This observational protocol, based on a make believe picnic game, was inspired by the Peanut Butter Protocol (Creaghead, 1984). The examiner follows a protocol whereby he invites the child, in a natural way, to express 23 communicative intentions or rules of communication. For example, the examiner may probe the communicative intention "request for action" by asking the child to open a bottle of juice with a cap that cannot be opened by children. The examiners are trained to follow the children's lead if the situation presents itself (e.g., if the children ask a question) in order to promote a natural conversation, while continuing to follow the protocol as they go along. Each of the first 50 utterances produced by the child is coded according to the presence or absence of criteria pertaining to the 14 variables of the PSCS-P, except for the variables "number of words per minute" and "number of utterances per minute," for which the numbers are tallied. The speech samples of this study were codified by the same person (the principal author) to increase the reliability of this measure. The results are then compiled into an *Excel* file and formulas are used to convert the results into a percentage of success.

Executive functions

Four neuropsychological tests were used to assess self-control, inhibition, flexibility, WM, and planning. Although these tests are not commercialized tools, they are frequently used in research in the absence of tests with better psychometric properties for preschoolers (Monette and Bigras, 2008).

The Prohibited Toy protocol was used to measure *self-control* ability (Rasmussen et al., 2008). This task correlates with other tests involving "hot" inhibition (Monette and Bigras, 2008), which refers to the cognitive process controlling decision-making that entails an emotional or motivational issue (Hongwanishkul et al., 2005; Zelazo and Müller, 2005). In the Prohibited Toy task, the examiner asks the child to turn his back so that they can play a guessing game. After two successful guesses (which animal corresponds to the sound made by a toy animal) the examiner announces to the child that he has to leave for a minute. Before leaving, the examiner asks the child not to look at the object behind him so that they may continue the guessing game upon the examiner's return. No points are awarded if the child looks at the object and one point is attributed if the child does not turn around to look.

The Backwards Digit Span (BDS) was used to assess *working memory* in an auditory-verbal modality (Davis and Pratt, 1995). In Davis and Pratt's protocol (1995), the examiner demonstrates

Table 1 | Description of the pragmatic skills coding system—preschool version.

CONVERSATIONAL COMPLEXITY SCALE ($\alpha = 0.68$)	
Turnabout	Percentage of utterances that have the dual function of responding to the interlocutor and restarting the conversation by adding information [e.g., "But" (response) "This glass will be mine." (expansion) (ICC = 0.67[a])].
Organization of utterances	Percentage of utterances that link more than one piece of information (regarding people, objects, time, location, action, etc.) in a single utterance [e.g., "I'm (subject) gonna eat (action) grapes (object)." (ICC = 0.79)].
Number of new themes	Percentage of utterances that produce new themes (ICC = 0.62).
Abstraction level of themes	Percentage of utterances that introduce themes that are decontextualized in time (e.g., I'm gonna go skiing this winter), place or reality (fictitious/fantasy) (e.g., You you're the mom and I'm the dad (ICC = 0.89).
TALKATIVENESS SCALE ($\alpha = 0.71$)	
Number of words	Number of words per minute (ICC = 1.00).
Number of utterances	Number of utterances per minute (ICC = 1.00).
Number of utterances per speaking turn	Percentage of utterances that express more than one utterance (separated by a delay of more than 2 s) per speaking turn (ICC = 0.93).
ASSERTIVENESS SCALE ($\alpha = 0.66$)	
Initiations	Percentage of utterances that initiate conversation, rather than answering a question (ICC = 0.88).
Requests	Percentage of utterances that formulate requests (ICC = 0.56).
Conversation breakdown repairs	Percentage of utterances that repair conversation breakdowns (e.g., child: "Box.," research assistant: "What?," child: "The box." (ICC = 0.52).
COMMUNICATIVE CONTROL SCALE ($\alpha = 0.38$)	
Fluidity	Percentage of utterances that are free of involuntary and unnecessary repetition or hesitation (e.g., "I want the. . . the bottle") (ICC = 0.93).
Non-interruption	Percentage of utterances that do not interrupt the interlocutor (ICC = 0.72).
RESPONSIVENESS SCALE ($\alpha = 0.61$)	
Contingency	Percentage of utterances that adequately respond to a request by the interlocutor (e.g., research assistant: "Will you play with the puzzle?," child: "OK.") (ICC = 0.81).
Utterance clarity	Percentage of utterances that express clear and understandable statement (ICC = 0.14[b]).

[a] ICC, Intraclass correlation coefficient. The speech samples of this study were codified by the same person. However, the principal author and an undergraduate student codify eighteen speech samples separately, during the validation process of the PSCS-P, in order to compute the ICC of each variable.

[b] This variable's ICC is below the "fair" level of 0.40 suggested by Cicchetti (1994). But when the inter-rater reliability is calculated in terms of percentage of agreement, the rate of this variable still remains relatively high at 91%, even higher than other variables. The lack of variability in this variable seems to have reduced the ICC.

to the child how to repeat a series of two numbers backwards using a puppet. The examiner then notes the longest series of numbers that the child manages to repeat backwards. The child is assigned a score of one if he fails to repeat two digits backwards, a score of two if he can recall two and so on.

The Dimensional Change Card Sort (DCCS) was used to measure *flexibility* (Zelazo, 2006). In this test, the examiner shows the child two target cards, a blue rabbit and a red boat, and asks the child to sort a set of cards, assigning each card either to the "red rabbit" pile or the "blue boat" pile. In the first phase, the child must sort the cards according to the shape of the objects on them. In the second phase, the child must sort the cards according to their colors. In the third phase, the child must alternate between sorting the cards by color and sorting them by shape. The child receives one point if he succeeds in the first phase, two for the second phase and three for the third phase.

The Tower of Hanoï (ToH) was used to measure *planning* and *inhibition* (Welsh et al., 1991). In this test, the child must move three rings of increasing size around on three pegs. The aim is to reach the final position with all the rings in descending order on the peg to the right. This must be done within the least number of moves while observing three rules: (1) not to put a larger ring on top of a smaller one, (2) to move the rings one at a time and (3) not to place the rings anywhere but on the pegs. The examiner explains the rules using an analogy—referring to the rings as a family of squirrels (i.e., smaller = child, medium = mother and larger = father)—and a demonstration. The examiner then makes sure the child understands the rules by asking him to perform the allowed moves. The child is entitled to six trials for each new problem. If he finds the solution within the designated number of moves on the first trial, he is assigned 6 points. One point is subtracted each time the child needs an additional trial to solve the problem within the designated number of moves. If the child fails to solve the problem within the designated number of moves after six trials, the examiner does not administer the following problems. The planning score is computed based on the total number of points, with a maximum score of 36 points (6 points for each of 6 problems). The inhibition score is computed by calculating the number of illegal moves over the total number of trials played (Ahonniska et al., 2000). The term "inhibition" is used here to differentiate it from the self-control measure evaluated by the Prohibited Toy protocol. The inhibition score on the ToH can be considered a cool type of inhibition because, as opposed to the Prohibited Toy protocol, the goal of the task is more cognitive and has no emotional underpinning (Hongwanishkul et al., 2005; Zelazo and Müller, 2005).

A principal component analysis (PCA) was performed on the EF measures to ensure that it was statistically possible to create a composite score with these measures. The flexibility score, however, was removed from the composite score because of its lack of interindividual variability (see **Table 2**) and the absence of any significant correlation with the other EF measures ($r =$ 0.06 to 0.22, $p > 0.05$). The PCA resulted in a one-factor solution, explaining 56.63% of the variance in the four remaining EF scores. Consequently, the composite score was computed by tallying the scores of each measures in standardized score.

Estimated intellectual quotient

The Peabody Picture Vocabulary Test—Revised (PPVT-R, French version) (Dunn et al., 1993) and the Block Design from the Wechsler Preschool and Primary Scale of Intelligence, 3rd edition (WPPSI-III) (Wechsler, 2002) were chosen to represent verbal (Fagan et al., 2007) and non-verbal IQ (Sattler, 2008). The PPVT-R evaluates *receptive vocabulary*. In this task, the child is presented with a set of four pictures. The examiner asks the child to point to the picture that corresponds to the word he says. The Block Design from the WPPSI-III was used to assess *visuoconstructive abilities*. In this test, the child is asked to reproduce several two-dimensional models with blocks, as fast as he can. The raw results of these tests were used for the purposes of analysis to facilitate comparison with the EF tests, for which normative data were not available.

A second PCA was performed on the measures used to estimate IQ, namely, vocabulary and visuoconstructive abilities, with the objective of creating another composite score. A one-factor solution emerged explaining 61.52% of the variance. Thus, the PCA supported the aggregation of the results for vocabulary and visuoconstructive abilities into an IQ composite score. Again, results were computed by adding the scores for each of the measures in standardized score.

PROCEDURE

Participants were recruited in the fall of 2008. The participating children were recruited through five publicly funded childcare centers in the Montreal region. Parental consent for the participants' participation in the research project was given following a request by email and phone. The instruments were administered by three psychology students who had received 15 h of training on the administration of the instruments. Each child was individually tested at his childcare center during two 45-min periods. The examiners administered the PPVT-R and the observational protocol of the PSCS-P on the first day of testing. On the second day, they administered, in the following order, the Block Design subtest (WPPSI-III), the DCCS, the Prohibited Toy protocol, the BDS and the ToH. The childcare provider and the participating children received a book to thank them for their participation.

RESULTS

Table 2 presents the descriptive results for all the measures: (1) PS, evaluated using the five scales of the PSCS-P (conversational complexity, talkativeness, assertiveness, communicative control and responsiveness), (2) EF, assessed through measures of self-control, inhibition, WM, flexibility and planning and (3) IQ, estimated based on measures of receptive vocabulary and visuoconstructive abilities. In order to determine the distribution of participants across these measures, their scores were divided into three categories: low, medium and high. It should be noted that the children's PS, EF, and IQ scores were generally fairly well-distributed across these different categories. However, 71% of the children were assigned a medium flexibility score on the DCCS,

Table 2 | Descriptive statistics for the executive function (EF), intellectual quotient (IQ), and pragmatics skills (PS) measures.

Constructs	Measures	Range	Min-max	Scores distribution (%)			Means	*SD*
				Low	Medium	High		
PS								
Conversational complexity	PSCS-P	0–4	0.08–1.47	33	34	33	0.77	0.34
Talkativeness	PSCS-P	0–3	0.34–1.63	33	34	33	0.91	0.32
Assertiveness	PSCS-P	0–3	0.45–2.56	33	34	33	1.51	0.52
Communicative control	PSCS-P	0–2	1.73–2.00	33	34	33	1.88	0.07
Responsiveness	PSCS-P	0–2	1.56–2.00	31	38	30	1.84	0.09
EF								
Self-control	Forbidden toy	0–1	0–1	66	–	34	0.61	0.49
Inhibition	Towers of hanoï[a]	0–1	0.14–0.92	31	36	30	0.52	0.23
Working memory	Backward digit span	1–5	1–4	44	31	24	1.82	0.83
Flexibility	DCCS	1–3	1–3	17	71	11	1.94	0.54
Planning	Towers of hanoï[b]	0–36	0–32	33	34	33	15.29	8.37
IQ ESTIMATE								
Vocabulary	PPVT-R (French version)	0–175	23–93	33	34	33	60.59	18.24
Visuoconstructive abilities	Block design (WPPSI-III)	0–40	18–32	37	23	40	24.08	3.03

Raw scores are presented.

[a] *Number of illegal moves over the total number of trials played.*

[b] *Problem resolution scores.*

which means that this measure showed very low interindividual variability.

Prior to all inferential statistics, transformations were made to the data to reduce the inconvenience caused by missing data when administering the EF tests. These missing data (4.5%) were replaced by an algorithm of Expectation Maximization (EM) by calculating the expected scores based on the results of the other EF scores. This method was chosen because the missing data were randomly distributed across the various measures [MCAR Chi2 (8) = 14.35, $p > 0.05$] (Ervin-Tripp, 1978). In addition, one subject had a multivariate extreme value, detected by calculating the Mahalanobis D2. This subject's results on the ToH were very abnormal and thus were replaced by an EM algorithm using the results of the other EF tests. Moreover, some of the variables of the PSCS-P were not normally distributed. Logarithmic transformations were performed to normalize the "breakdown repairs," "non-interruption," "contingency," and "utterance clarity" variables. The "abstraction level of themes" variable was dichotomized based on the presence or absence of at least one decontextualized theme during the exchange.

Before addressing the main objective of this study, Pearson correlations performed in order to present the link between the sociodemographics characteristics, namely, age, gender, household income and education of the mother, and our measurements. These correlations, presented in **Table 3**, show that mother education has the strongest relation with children performance on the measure of PS, EF, and IQ (ranging from $r = -0.10$, $p > 0.05$ to $r = 0.32$, $p < 0.01$). Both age and income correlate significantly with vocabulary (respectively $r = 0.26$, $p < 0.05$ and $r = 0.36$, $p < 0.01$) and planning (respectively $r = 0.33$, $p < 0.01$ and $r = 0.30$, $p < 0.05$) for instance. On the other hand, gender is only significantly associated with talkativeness ($r = 0.27$, $p < 0.05$), indicating that boys are more talkative than girls.

Table 3 | Pearson correlations between sociodemographic characteristics and executive functions (EF), intellectual quotient (IQ), and pragmatics skills (PS) measures.

	Age (month)	Gender	Income	Mother's education
PS				
Complexity	−0.07	0.17	−0.05	0.31*
Talkativeness	−0.14	0.27*	−0.18	0.15
Assertiveness	−0.22	0.05	0.04	0.28*
Communicative control	0.08	0.07	−0.02	−0.10
Responsiveness	0.17	0.18	−0.19	−0.07
EF				
Self-control	0.02	−0.15	−0.02	−0.09
Inhibition	0.23	−0.10	0.25*	0.14
Working memory	0.12	0.06	0.05	0.27*
Flexibility	0.32**	0.10	0.24	0.32**
Planning	0.33**	−0.02	0.30*	0.30*
IQ ESTIMATED				
Vocabulary	0.26*	0.15	0.36**	0.19
Visuoconstructive	0.15	0.22	0.02	0.28*

*$p < 0.05$, **$p < 0.01$.

As for the inferential statistics, **Table 4** presents the Pearson correlations performed to determine what role self-control, inhibition, flexibility, WM, planning and the EF composite score (sum of all EF measures except flexibility) played in the children's PS. In order to determine whether the contribution of EF to PS was significantly different from that of IQ to PS, differences among the correlation coefficients were tested using the Fisher z transformation formula proposed by Meng et al. (1992). On the whole, these analyses showed that EF correlated with PS differently than IQ for 2 of the 5 scales in the PSCS-P and 3 of the 14 associated variables (see **Table 4**).

These correlation results are presented in more detail according to each of the five categories of PS: conversational complexity, talkativeness, assertiveness, communicative control and responsiveness. With respect to conversational complexity, no relationship between EF and PS was strong enough to reach the significance threshold. However, the conversational complexity scale ($z = 2.10$, $p < 0.05$) and its variable related to the level of organization of the information in the utterances ($z = 2.38$, $p < 0.05$) correlated significantly differently with EF than with IQ. Specifically, the EF correlation showed a negative tendency with regard to these PS, whereas the IQ correlation showed a positive tendency. Although the EF and IQ correlations with these PS were not significant, the fact that they went in opposite directions resulted in a significant difference.

Regarding talkativeness, both the EF composite score and inhibition were associated with a decrease in the talkativeness scale ($r = -0.24$ and -0.28, $p < 0.05$). They were also related to a decrease in the variable of this scale measuring the number of utterances per speaking turn ($r = -0.28$, $p < 0.05$ and $r = -0.40$. $p < 0.01$). Moreover, self-control was related to a reduction in the number of words per minute, at a marginally significant level ($r = -0.24$, $p < 0.06$). For talkativeness ($z = 2.04$, $p < 0.05$) and number of utterances per speaking turn only ($z = 3.02$, $p < 0.01$), the strength of the EF correlation coefficients differed significantly from the strength of the IQ correlation coefficients. In fact, IQ was related to an increase in these three PS, but did not make a significant contribution to them.

Furthermore, assertiveness yielded a similar correlation pattern to talkativeness and conversational complexity. Again, EF showed a more negative tendency, whereas IQ showed a more positive correlation with PS in general. WM was correlated significantly with a reduction in the number of requests ($r = -0.25$, $p < 0.05$). Three other marginally significant relationships involving EF were also found, all of them being negative. One of these relationships showed that the EF composite score was correlated with the assertiveness scale ($r = -0.23$, $p < 0.06$). The other two showed that self-control was related to a reduction in the number of communication breakdown repairs ($r = -0.24$, $p < 0.06$) and a decrease in the assertiveness scale in general ($r = -0.23$, $p < 0.06$). Although none of the predictors were significantly correlated with the capacity to initiate conversation, the correlation coefficients for EF ($r = -0.22$, $p > 0.05$) and IQ ($r = 0.09$, $p > 0.05$) were significantly different from one another ($z = 2.17$, $p < 0.05$). Once again, the difference in the direction

Table 4 | Pearson correlations between pragmatic skills (PS) and executive functions (EF) and between PS and intellectual quotient (IQ); and results of the test of differences between the correlation coefficients for the two relationships.

PS	EF					IQ		EF[a]	IQ[b]	rPSxEF ≠ rPSxIQ[c]	
Scales variables	**Self-control**	**Inhibition**	**WM**	**Flexibility**	**Planning**	**Vocabulary**	**VC**			***Z***	**(*p*)**
Conver. complexity	−0.20	−0.09	0.03	0.20	−0.10	0.09	0.18	−0.12	0.18	2.10	0.04*
Turnabout	−0.12	−0.07	0.03	0.20	−0.10	0.05	0.16	−0.09	0.14	1.61	0.11
Organization of utterances	−0.18	−0.15	0.06	0.10	−0.05	0.17	0.20	−0.11	0.23[t]	2.38	0.02*
Number of new themes	−0.20	−0.07	0.01	0.16	−0.06	0.12	0.10	−0.11	0.14	1.75	0.08
Abstraction level of themes	−0.09	0.03	0.04	0.14	−0.11	−0.08	0.12	−0.04	0.03	0.63	0.53
Talkativeness	−0.19	−0.28*	−0.09	0.04	−0.15	−0.03	0.10	−0.24*	0.05	2.04	0.04*
Number of words	−0.24[t]	−0.18	−0.05	0.07	−0.15	0.05	0.03	−0.21	0.05	1.83	0.07
Number of utterances	−0.14	−0.09	−0.02	−0.04	−0.04	−0.17	0.02	−0.10	−0.09	0.07	0.94
Utterances per speaking turn	−0.10	−0.40**	−0.15	0.07	−0.18	0.05	0.18	−0.28*	0.15	3.02	0.002**
Assertiveness	−0.23[t]	−0.20	−0.17	0.14	−0.08	−0.01	0.02	−0.23[t]	0.01	1.69	0.09
Initiations	−0.19	−0.20	−0.12	0.16	−0.15	0.14	0.01	−0.22	0.09	2.17	0.03*
Requests	−0.12	−0.22	−0.25*	0.06	−0.04	−0.12	−0.02	−0.21	−0.09	0.85	0.39
Breakdown repairs	−0.24[t]	−0.05	−0.03	0.11	−0.01	−0.05	0.05	−0.11	−01	0.70	0.48
Communicative control	0.21	0.15	0.13	−0.01	0.26*	0.02	0.01	0.25*	0.02	1.63	0.10
Fluidity	0.30*	0.30*	0.25*	−0.06	0.31**	0.17	0.08	0.38**	0.15	1.69	0.09
Non-interruption	0.02	−0.07	−0.05	0.05	0.09	−0.14	−0.06	−0.01	0.13	0.98	0.33
Responsiveness	−0.04	0.00	0.29*	0.04	0.00	−0.06	0.09	0.08	0.02	0.42	0.67
Contingency	−0.06	0.03	0.25*	0.05	−0.04	−0.06	−0.01	0.07	−04	0.77	0.44
Utterance clarity	−0.02	−0.03	0.26*	0.01	0.03	−0.05	0.17	0.08	0.08	0	1.00

VC, visuoconstructive abilities; Conver. Complexity, conversational complexity.

[a] Flexibility was not included in the EF composite score.

[b] IQ was estimated using measures of receptive vocabulary and visuoconstructive abilities.

[c] Probability that the correlation between EF and PS is significantly different ($p < 0.05$) from that between IQ and PS using the Meng et al. (1992) method.

*[t] Marginally significant at $p < 0.06$, *$p < 0.05$, and **$p < 0.01$.*

of the correlation, EF being negative and IQ being positive, helped produce a significantly different correlation coefficient between the two predictors.

This difference in the direction of the predictor's relationship with PS was not observed for the communicative control and responsiveness scales. In fact, both EF and IQ tended to correlate positively with these PS and no correlation coefficient differed significantly. As regards communicative control, the most striking result was certainly that all of the measures included in the EF composite score were correlated with utterance fluidity ($r = 0.25$, $p < 0.05$ to $r = 0.31$, $p < 0.01$).

As for responsiveness, WM was positively correlated with the responsiveness scale ($r = 0.29$, $p < 0.05$) and its two variables, namely, contingency ($r = 0.25$, $p < 0.05$) and utterance clarity ($r = 0.26$, $p < 0.05$). No other predictor was correlated with the responsiveness scale or its variables.

It should be noted that no significant correlations were found between PS and the IQ composite score, or the variables on which it was based, namely, vocabulary and visuoconstructive abilities. Nevertheless, there was a marginally significant relationship between IQ and the level of organization of the information in the utterances, a variable associated with the conversational complexity scale.

Table 5 | Summary of standard multiple regression analysis for the executive functions processes predicting utterance fluidity.

	B	***SE B***	β	***p***
Constant	0.87	0.02	–	
Self-control	0.03	0.02	0.20	0.15
Inhibition	0.02	0.04	0.06	0.72
Working memory	0.04	0.04	0.15	0.23
Planning	0.03	0.04	0.13	0.43

Given that more than one EF process correlated with utterance fluidity, a standard multiple regression analysis was performed between utterance fluidity (VD) and self-control, inhibition, WM and planning (VI) to calculate the total percentage of explained variance. The four VIs explained 15.2% of the variance associated with utterance fluidity [$F_{(4, 69)} = 2.91$, $p < 0.05$]. **Table 5** presents the beta coefficients for each individual predictor, none of which made a unique contribution to utterance fluidity. In others words, if the other predictors were held constant, none of these EF processes would contribute significantly to utterance fluidity.

Additionally, partial Pearson correlations were performed in order to control for the sociodemographics characteristics in the relationship between PS, EF, and IQ. **Table 6** presents Pearson correlations without others variables accounted for, and the partial Pearson correlations controlling, respectively, for age, gender, income, and education of the mother. Overall, results show little change in the significant level of the correlation after the control of the sociodemographics characteristics (those changes are highlighted in **Table 6**). In few instance the correlations became non-significants. Those instances involve for the most part the correlations implicating WM when controlling for age, gender, or income. It important to note that age, gender, and income did not make a significant contribution to WM (see **Table 3**) and therefore, the control of those variables seems to have introduced noise in the model. In others cases, mostly relating to the control of the education of the mother, the correlation significance level was raise.

DISCUSSION

The objective of this study was to further our understanding of the role of EF in the PS displayed by normally developing children while conversing with an adult. EF are generally defined as processes involved in new and complex tasks (Lezak et al., 2012), as are often social interaction where children most deploy their PS. For example, children between 5 and 7 years are likely to emit more than ten verbal and non-verbal behaviors to integrate a group of peers during play time (Dodge et al., 1986). To express these behaviors in a socially appropriate way, it seems logical to believe that EF like the capacity to anticipate the reactions of other, to plan behavior ahead, to adjust it along the way and to inhibit inappropriate behavior are involved.

Our results show, for instance that higher inhibitory control is associated with a decrease in talkativeness. This result could, at first glance, appear to be counter intuitive since EF should logically assist children with their PS rather than being detrimental to them. Nevertheless, our data are consistent with findings showing an excessive increase in talkativeness among individuals who likely have an inhibition deficit, such as children with ADHD (Landau and Milich, 1988; Humphries et al., 1994; Bruce et al., 2006) and patients with frontal lesions (Bernicot and Dardier, 2001). Arbuckle et al. (2000) revealed a more direct link between low inhibitory control and the tendency to provide more redundant information and be more talkative (marginally significant) among older adults (63–95 years) in a referential communication task. These authors alleged that poorer inhibitory skills could be associated with the intrusion of unnecessary information. It is a well-known finding that inhibitory control is needed to refrain from committing an intrusion error, for example, by retrieving the wrong word in a memory task (Levy and Anderson, 2002). In our study, children who made a greater number of illegal moves in the ToH task had a tendency to produce more than one utterance per speaking turn. This result was one of the more substantial effects found, as approximately 16% of the explained variance in the number of utterances per speaking turn could be accounted for by inhibitory control. It may be that children with higher inhibition skills are better at refraining from speaking more than is necessary, in this case, producing more than one utterance before their interlocutor started to speak again. In this sense, the rules of communication (e.g., respecting speaking turns) may act like the rules of a neuropsychological test such as the ToH. Between the age 2 and 4, Pellegrini et al. demonstrated that children tend to violate less frequently Gricean principles stipulating, for instance that an intervention should bring enough information, but not more than necessary (Grice, 1975). Thus, the decreased in talkativeness might perhaps indicate an increasing in the ability to follow this principal.

Another result was even more unexpected. Indeed, our data show marginally significant correlations between higher self-control ("hot" inhibition) and a decrease in the assertiveness scale and a reduction in the number of communication breakdown repairs. This was also a counter intuitive result since our measure of assertiveness was constructed as a positive concept. Notwithstanding, this result could be consistent with data showing that a lack of inhibition may lead to aggressive behavior (Raaijmakers et al., 2008), which could be viewed as a rare and high amplitude subclass of assertive behavior (Patterson et al., 1967; Ostrov et al., 2006). Of course, correcting the interlocutor's miscomprehension does not correspond to an aggressive behavior because it does not harm this person in any way.

Yet, it is important to recall that the participants in our study were asked to interact with a research assistant with whom they were unfamiliar. Typically, children are much more reserved with an adult with whom they are not acquainted, which may tend to reduce their overall level of assertiveness. In fact, Bishop et al. (1994) showed that, compared to children with Semantic-Pragmatic Disorder, normally developing children had a slightly greater tendency (although not significant, $p = 0.09$) to initiate conversation with a familiar adult than with an unfamiliar one. This means that a low degree of assertiveness with an unfamiliar adult could be a sign of better PS, meaning that the child is able to adapt to the context of the situation. In our observational protocol, the examiner asked the child what color of grapes he wanted and then gave him the other color on purpose. This procedure was used to see whether the child would repair the communication breakdown. As said previously, children with better self-control tended to refrain from correcting the research assistant. If we consider the perspective of a 4 year-old child meeting an unfamiliar adult, it is easy to see why the child might be intimidated by the adult and refrain from correcting him. On the other hand, a child with low self-control may be more inclined to act the same way in any situation, and thus be more likely to correct the research assistant as he would do with a friend. Consequently, self-control may help children refrain from overly asserting themselves when the situation precludes it. Also, we did not take into account the manner used to correct the adult. Future research is needed to evaluate the relationship between the quality of assertiveness and EF, as opposed to the quantity measure used in our study.

Moreover, the above-mentioned negative correlations between inhibition and talkativeness and between self-control and assertiveness lead us to question the linear design of these scales, which presume that a higher score is always better. It may instead be that the ideal level of talkativeness and assertiveness is moderate, neither too high nor too low. Thus, the child should try to adapt to his interlocutor by speaking about the same amount as

Table 6 | Partial Pearson correlations between pragmatics skills (PS) and executive functions (EF) and intellectual quotient (IQ) controlling for age, gender, income, and education of the mother.

PS	Variables control for	EF					IQ	
Scales variables		**SC**	**Inhi**	**WM**	**Flex**	**Plan**	**Voca**	**VC**
Complexity	No	−0.20	−0.09	0.03	0.20	−0.10	0.09	0.18
	Age	−0.15	−0.05	0.09	0.21	−0.11	0.08	0.21
	Gender	−0.13	−0.04	0.06	0.20	−0.11	0.05	0.18
	Income	−0.13	0.01	0.13	0.26*	−0.06	0.13	0.19
	Educ.	−0.13	−0.07	0.00	0.19	−0.19	0.03	0.12
Turnabout	No	−0.12	−0.07	0.03	0.20	−0.10	0.05	0.16
	Age	−0.09	−0.02	0.09	0.22	−0.07	0.07	0.19
	Gender	−0.07	−0.04	0.03	0.19	−0.10	0.02	0.14
	Income	−0.09	0.01	0.12	0.26*	−0.07	0.14	0.19
	Educ.	−0.07	−0.07	−0.03	0.20	−0.18	0.02	0.10
Organization of utterances	No	−0.18	−0.15	0.06	0.10	−0.05	0.17	0.20
	Age	−0.13	−0.10	0.12	0.12	−0.01	0.19	0.22
	Gender	−0.12	−0.11	0.08	0.10	−0.04	0.15	0.19
	Income	−0.09	−0.01	0.14	0.19	0.07	0.22	0.16
	Educ.	−0.11	−0.13	0.00	0.10	−0.14	0.13	0.11
Number of new themes	No	−0.20	−0.07	0.01	0.16	−0.06	0.12	0.10
	Age	−0.16	−0.07	0.03	0.17	−0.10	0.09	0.11
	Gender	−0.16	−0.06	0.05	0.17	−0.07	0.10	0.13
	Income	−0.15	−0.04	0.05	0.22	−0.09	0.09	0.10
	Educ.	−0.14	−0.06	−0.03	0.17	−0.15	0.07	0.03
Abstraction level of themes	No	−0.09	0.03	0.04	0.14	−0.11	−0.08	0.12
	Age	−0.07	0.04	0.02	0.13	−0.13	−0.11	0.10
	Gender	−0.05	0.06	0.02	0.13	−0.10	−0.11	0.08
	Income	−0.08	0.07	0.10	0.11	−0.09	−0.05	0.14
	Educ.	−0.07	0.04	0.04	0.10	−0.12	−0.11	0.12
Talkativeness	No	−0.19	−0.28*	−0.09	0.04	−0.15	−0.03	0.10
	Age	−0.17	−0.25*	−0.04	0.06	−0.12	−0.01	0.12
	Gender	−0.14	−0.25*	−0.11	0.03	−0.16	−0.08	0.05
	Income	−0.18	−0.20	0.02	0.14	−0.10	0.10	0.13
	Educ.	−0.16	−0.27*	−0.12	0.09	−0.18	−0.03	0.07
Number of words	No	−0.24	−0.18	−0.05	0.07	−0.15	0.05	0.03
	Age	−0.21	−0.14	0.01	0.09	−0.12	0.07	0.06
	Gender	−0.19	−0.16	−0.06	0.06	−0.15	0.01	−0.01
	Income	−0.22	−0.10	0.04	0.18	−0.08	0.19	0.05
	Educ.	−0.21	−0.19	−0.11	0.11	−0.21	0.05	−0.02
Number of utterance	No	−0.14	−0.09	−0.02	−0.04	−0.04	−0.17	0.02
	Age	−0.14	−0.07	0.02	−0.02	−0.01	−0.16	0.04
	Gender	−0.11	−0.08	−0.04	−0.05	−0.04	−0.21	−0.03
	Income	−0.17	−0.03	0.07	0.09	0.04	−0.03	0.07
	Educ.	−0.15	−0.12	−0.09	0.03	−0.07	−0.16	0.00
Utterances per speaking turn	No	−0.10	−0.40**	−0.15	0.07	−0.18	0.05	0.18
	Age	−0.07	−0.38**	−0.12	0.08	−0.17	0.06	0.20
	Gender	−0.05	−0.38**	−0.16	0.06	−0.18	0.01	0.15
	Income	−0.05	−0.37**	−0.06	0.07	−0.19	0.09	0.20
	Educ.	−0.04	−0.36**	−0.10	0.08	−0.17	0.03	0.19

(Continued)

Table 6 | Continued

PS	Variables control for	EF					IQ	
Scales variables		**SC**	**Inhi**	**WM**	**Flex**	**Plan**	**Voca**	**VC**
Assertiveness	No	−0.23	−0.20	−0.17	0.14	−0.08	−0.01	0.02
	Age	−0.20	−0.15	−0.08	0.17	−0.03	0.02	0.06
	Gender	−0.19	−0.19	−0.14	0.14	−0.09	−0.04	0.03
	Income	−0.21	−0.19	−0.16	0.20	−0.12	−0.03	0.02
	Educ.	−0.20	−0.24	−0.27*	0.14	−0.20	−0.08	−0.06
Initiation	No	−0.19	−0.20	−0.12	0.16	−0.15	0.14	0.01
	Age	−0.14	−0.17	−0.07	0.18	−0.14	0.14	0.02
	Gender	−0.14	−0.18	−0.09	0.17	−0.16	0.12	0.01
	Income	−0.14	−0.16	−0.06	0.21	−0.14	0.18	0.01
	Educ.	−0.14	−0.20	−0.15	0.17	−0.21	0.11	−0.04
Requests	No	−0.12	−0.22	−0.25*	0.06	−0.04	−0.12	−0.02
	Age	−0.10	−0.18	−0.16	0.10	0.04	−0.08	0.03
	Gender	−0.10	−0.23	−0.22	0.07	−0.05	−0.13	0.01
	Income	−0.14	−0.21	−0.20	0.10	−0.07	−0.09	0.01
	Educ.	−0.11	−0.25*	−0.31*	0.07	−0.11	−0.18	−0.05
Breakdown repairs	No	−0.24	−0.05	−0.03	0.11	−0.01	−0.05	0.05
	Age	−0.22	−0.01	0.04	0.13	0.03	−0.02	0.09
	Gender	−0.21	−0.04	−0.02	0.10	−0.02	−0.07	0.05
	Income	−0.23	−0.09	−0.12	0.16	−0.08	−0.17	0.03
	Educ.	−0.23	−0.10	−0.17	0.08	−0.15	−0.12	−0.05
Communicative control	No	0.21	0.15	0.13	−0.01	0.26*	0.02	0.01
	Age	0.18	0.14	0.09	−0.02	0.28*	0.02	−0.01
	Gender	0.20	0.16	0.08	−0.02	0.28*	0.02	−0.03
	Income	0.20	0.16	0.12	−0.05	0.30*	0.01	−0.02
	Educ.	0.20	0.17	0.15	−0.04	0.32**	0.03	0.00
Fluidity	No	0.30*	0.30*	0.25*	−0.06	0.31**	0.17	0.08
	Age	0.28*	0.29*	0.23	−0.07	0.33**	0.19	0.06
	Gender	0.29*	0.30*	0.22	−0.07	0.32**	0.18	0.05
	Income	0.32*	0.32*	0.20	−0.04	0.39**	0.17	0.04
	Educ.	0.29*	0.32*	0.26*	−0.04	0.37**	0.22	0.08
Non-interruption	No	0.02	−0.07	−0.05	0.05	0.09	−0.14	−0.06
	Age	−0.01	−0.09	−0.10	0.04	0.08	−0.15	−0.08
	Gender	0.00	−0.07	−0.09	0.04	0.10	−0.14	−0.10
	Income	−0.01	−0.08	−0.02	−0.04	0.06	−0.17	−0.07
	Educ.	0.01	−0.06	−0.04	−0.02	0.11	−0.18	−0.08
Responsiveness	No	−0.04	0.00	0.29*	0.04	0.00	−0.06	0.09
	Age	−0.04	−0.04	0.27*	0.02	−0.07	−0.12	0.07
	Gender	−0.02	0.02	0.29*	0.03	0.00	−0.09	0.06
	Income	−0.04	0.08	0.37**	0.10	0.05	−0.01	0.08
	Educ.	−0.06	0.03	0.36**	0.12	0.00	−0.05	0.12
Contingency	No	−0.06	0.03	0.25*	0.05	−0.04	−0.06	−0.01
	Age	−0.05	−0.01	0.26*	0.03	−0.11	−0.12	−0.01
	Gender	−0.03	0.04	0.28*	0.04	−0.05	−0.10	−0.03
	Income	−0.07	0.08	0.36**	0.10	−0.03	−0.01	0.01
	Educ.	−0.07	0.04	0.33**	0.11	−0.06	−0.07	0.02

(Continued)

Table 6 | Continued

PS	Variables control for	EF					IQ	
Scales variables		**SC**	**Inhi**	**WM**	**Flex**	**Plan**	**Voca**	**VC**
Utterance clarity	No	−0.02	−0.03	0.26*	0.01	0.03	−0.05	0.17
	Age	−0.02	−0.06	0.22	−0.01	−0.01	−0.09	0.14
	Gender	0.00	−0.01	0.24	0.00	0.05	−0.07	0.13
	Income	−0.01	0.05	0.23*	0.08	0.12	0.00	0.14
	Educ.	−0.03	0.01	0.30*	0.10	0.07	−0.02	0.19

SC, sefl-control; Inhi, inhibition; WM, working memory; Flex, flexibility; Plan planning; Voca, vocabulary; VC, vioconstructive abilities; Educ., education of the mother. The shaded cells indicate a change in the level of significance in the correlations between PS, EF, or IQ after controlling the sociodemographic characteristics.
$^{*}p < 0.05$, $^{**}p < 0.01$.

the latter and acting more thoughtfully, and the child's inhibition level may help him to achieve this.

Furthermore, one of the most impressive findings of this study is the involvement of all EF measures (except flexibility[1]) in the production of more fluid utterances. These results corroborate those of Engelhardt et al. (2013) showing that inhibition was linked to a decrease in dysfluencies among adolescents and adults in a sentence production task. According to their study and previous others (Berg and Schade, 1992; Dell et al., 1997; Engelhardt et al., 2010), inhibition may help reduce the risk of articulating the wrong word by inhibiting the competing phrasing. Our data confirm the entanglement of both "hot" (i.e., emotional) and "cool" (i.e., cognitive) types of inhibition in utterance fluidity in a more natural setting. They also suggest the involvement of WM and planning. On the other hand, vocabulary and visuoconstructive abilities did contribute significantly to the articulation of fluid utterances. Yet, their correlations were not significantly different from those between EF and utterance fluidity, meaning that their role is not much different.

Moreover, the children in our study with a high WM capacity were more likely to formulate contingent answers and produce utterances that could be clearly understood by the interlocutor. They also had a tendency to make fewer requests. The WM or its verbal counterpart, phonological short-term memory, has long been suspected to be involved in language comprehension and production in general (Bock, 1982; Gathercole and Baddeley, 1990; Just and Carpenter, 1992). It has been proposed that the primary function of phonological short-term memory may be to support the long-term learning of the phonological structure of language (Baddeley et al., 1998; Gathercole et al., 2005). There is evidence of this theory in others studies involving speech samples from young children. Indeed, phonological short-term memory in 3–4 year old children has been linked to their ability to formulate more complex utterances in terms of the structural aspects of language, such as the number of words, syntax and vocabulary variety (Adams and Gathercole, 1995; Gathercole, 2000). Although fewer studies have focused on the social aspect of language, there is nevertheless data demonstrating the involvement of WM in PS. For instance, WM has been associated with the interpretation of irony among normally developing children aged 5–9 years (Filippova and Astington, 2008). This increase in language comprehension, and even social understanding, may help children better grasp the situation at hand and consequently respond in a more socially appropriate way.

The negative correlation between WM and the number of requests was more surprising, since requests are sometimes viewed as a more complex communicative intention (Favre and Maeder, 2002). According to our qualitative observations while coding the children's utterances, many of their requests had to do with comprehension (e.g., "What?"). As previously stated, WM is essential for language comprehension. In this sense, children with lower WM may have had more difficulty understanding the interlocutor's speech than other children and may therefore have asked more questions to improve their comprehension. Future studies are needed to confirm this interpretation, especially since we did not measure which types of requests WM was related to.

In sum, EF appear to help preschool children better filter speech, control their level of assertiveness, refrain from articulating utterances incorrectly and respond in a socially appropriate way. Verbal and non-verbal cognitive abilities appear to offer a small, but positive contribution to PS. The effect of EF, on the other hand, appears to be greater than and not always in the same direction as that of IQ. Therefore, EF processes appear to affect PS in a unique and specific way, separately from the more global affect driven by cognitive maturation. Overall, our results suggest that EF play a more important role than IQ in the PS exhibited by children in a semi-structured conversational setting. Indeed, receptive vocabulary and visuoconstructive skills, which were combined to estimate IQ, did not make a significant contribution to any PS. Perhaps the new and unpredictable characteristics of live social interaction are more likely to involve EF.

LIMITATIONS

It is important to note that our results indicate that the influence of EF and IQ on pragmatic skills is generally limited. This means that a large part of the variance can still be accounted for by other factors such as the child's temperament (Coplan and Weeks, 2009) or socialization experiences (Bruner, 2002). Yet, some studies have found a much larger effect size between EF and PS. Douglas (2010), for instance, reported that, among adults with severe TBI, as high as 37% of PS variation (evaluated using

[1]This absence of a significant relationship may be caused by a lack of interindividual variability, as previously stated.

the La Trobe Communication Questionnaire) could be explained by executive functioning. In comparison, the strongest relationships found in our study were approximately 15% of explained variance, less than half of the effect size found by Douglas (2010). It could be argued that our sample was composed of a relatively homogeneous group of children (all aged between 3;10 and 5;7 years, typically developing, attending childcare in the same area and mostly raised by educated mothers). This homogeneity may have reduced the variability in our measures and thus the strength of the correlations we were able to obtain.

The lack of interindividual variability seems to have predominantly affected the ability to measure the relationship between flexibility, as measured using the DCCS, and the other variables. Indeed, the flexibility score could not discriminate between the children in our sample (71% of the children had the same score) and did not correlate significantly with the other variables in our study. It would therefore be pertinent in the future to use more sensitive measures of flexibility to differentiate between different levels of cognitive flexibility among 4–5 year old children. Monette and Bigras (2008) have suggested that Hughes' (1998) set-shifting task and the Trail Making Test for preschoolers (Espy and Cwik, 2004) could serve as alternatives to the DCCS, particularly for typically developing children in this age group. Future research could alternatively use a more widespread aged group to increase the interindividual variability of this measure.

It is also necessary to recall the exploratory nature of this study. A large number of statistical analyzes were performed, which has the effect of increasing the probability of a family-wise error rate. Further studies are needed to replicate these results, especially since this is the first study to have used the PSCS-P to examine the link between PS and EF. In addition, the EF tests used in this study did not come from commercialized tools since few such tools are available for preschoolers (Monette and Bigras, 2008). More research should be conducted to develop and validate EF measures for children in this age group. We should also specify that our results came from a single measurement time, thus making it impossible to study the effect of EF on the development of pragmatics. A longitudinal study using multiple time points would make it possible to examine the cognitive factors underlying the acquisition of pragmatics.

CONCLUSION

To conclude, research into the cognitive factors that contribute to the acquisition of pragmatics among children is in the beginning stages. Further research involving normally developing children is needed in order to better understand how children acquire pragmatic skills, an ability that is essential to their social development and academic achievement (Ervin-Tripp, 1978; Black and Hazen, 1990; Lemelin and Boivin, 2007; McKown, 2007; Coplan and Weeks, 2009; Brinkman et al., 2013).

ACKNOWLEDGMENTS

This research was supported by grants awarded by the CRSH (Canada), FQRSC (Canada), and the PAFARC (UQAM, Canada). We thank the children and the childcare center for their participation in the study.

REFERENCES

Adams, A. M., and Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *J. Speech Hear. Res.* 38, 403.

Adams, C. (2002). Practitioner review: the assessment of language pragmatics. *J. Child Psychol. Psychiatry* 43, 973–987. doi: 10.1111/1469-7610.00226

Ahonniska, J., Ahonen, T., Aro, T., and Lyytinen, H. (2000). Suggestions for revised scoring of the tower of hanoi test. *Assessment* 7, 311–319. doi: 10.1177/107319110000700309

Arbuckle, T. Y., Nohara-Leclair, M., and Pushkar, D. (2000). Effect of off-target verbosity on communication efficiency in a referential communication task. *Psychol. Aging* 15, 65. doi: 10.1037/0882-7974.15.1.65

Baddeley, A., Gathercole, S., and Papagno, C. (1998). The phonological loop as a language learning device. *Psychol. Rev.* 105, 158–173. doi: 10.1037/0033-295X.105.1.158

Berg, T., and Schade, U. (1992). The role of inhibition in a spreading-activation model of language production. I. The psycholinguistic perspective. *J. Psychol. Res.* 21, 405–434.

Bernicot, J., and Dardier, V. (2001). Communication deficits: assessment of subjects with frontal lobe damage in an interview setting. *Int. J. Lang. Commun. Disord.* 36, 245–263. doi: 10.1080/13682820010019937

Bishop, D., Hartley, J., and Weir, F. (1994). Why and when do some language-impaired children seem talkative? A study of initiation in conversations of children with semantic-pragmatic disorder. *J. Autism Dev. Disord.* 24, 177–197. doi: 10.1007/BF02172095

Bishop, D. V. M., and Adams, C. (1991). What do referential communication tasks measure? A study of children with specific language impairment. *Appl. Psycholinguist.* 12, 199–215. doi: 10.1017/S0142716400009140

Bishop, D. V. M., and Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology I: generativity. *Autism* 9, 7–27. doi: 10.1177/1362361305049027

Black, B., and Hazen, N. L. (1990). Social status and patterns of communication in acquainted and unacquainted preschool children. *Dev. Psychol.* 26, 379–387. doi: 10.1037/0012-1649.26.3.379

Blair, C., Zelazo, P. D., and Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Dev. Neuropsychol.* 28, 561–571. doi: 10.1207/s15326942dn2802_1

Bock, J. K. (1982). Toward a cognitive psychology of syntax: information processing contributions to sentence formulation. *Psychol. Rev.* 89, 1–47. doi: 10.1037/0033-295X.89.1.1

Bonifacio, S., Girolametto, L., Bulligan, M., Callegari, M., Vignola, S., and Zocconi, E. (2007). Assertive and responsive conversational skills of Italian-speaking late talkers. *Int. J. Lang. Commun. Disord.* 42, 607–623. doi: 10.1080/13682820601084386

Brinkman, S., Gregory, T., Harris, J., Hart, B., Blackmore, S., and Janus, M. (2013). Associations between the early development instrument at age 5, and reading and numeracy skills at ages 8, 10 and 12: a prospective linked data study. *Child Indic. Res.* 6, 695–708. doi: 10.1007/s12187-013-9189-3

Bruce, B., Thernlund, G., and Nettelbladt, U. (2006). ADHD and language impairment: a study of the parent questionnaire FTF (Five to Fifteen). *Eur. Child Adoles. Psychiatry* 15, 52–60. doi: 10.1007/s00787-006-0508-9

Bruner, J. S. (2002). *Comment les Enfants Apprennent à Parler.* Paris: Retz.

Carpenter, M., Nagell, K., and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, 176. doi: 10.2307/1166214

Channon, S., and Watts, M. (2003). Pragmatic language interpretation after closed head injury: Relationship to executive functioning. *Cogn. Neuropsychiatry* 8, 243–260. doi: 10.1080/135468000344000002

Coplan, R. J., and Weeks, M. (2009). Shy and soft-spoken: shyness, pragmatic language, and socio-emotional adjustment in early childhood. *Infant Child Dev.* 18, 238–254. doi: 10.1002/icd.622

Creaghead, N. (1984). Strategies for evaluating and targeting pragmatic behaviors in young children. *Semin. Speech Lang.* 5, 241–252. doi: 10.1055/s-0028-1085181

Dardier, V., Bernicot, J., Delanoë, A., Vanberten, M., Fayada, C., Chevignard, M., et al. (2011). Severe traumatic brain injury, frontal lesions, and social aspects of language use: a study of French-speaking adults. *J. Commun. Disord.* 44, 359–378. doi: 10.1016/j.jcomdis.2011.02.001

Davis, H. L., and Pratt, C. (1995). The development of children's theory of mind: the working memory explanation. *Aust. J. Psychol.* 47, 25–31. doi: 10.1080/00049539508258765

Dell, G. S., Burger, L. K., and Svec, W. R. (1997). Language production and serial order: a functional analysis and a model. *Psychol. Rev.* 104, 123. doi: 10.1037/0033-295X.104.1.123

Dodge, K. A., Pettit, G. S., McClaskey, C. L., Brown, M. M., and Gottman, J. M. (1986). Social competence in children. *Monogr. Soc. Res. Child Dev.* 51, i-85. doi: 10.2307/1165906

Douglas, J. M. (2010). Relation of executive functioning to pragmatic outcome following severe traumatic brain injury. *J. Speech Lang. Hear. Res.* 53, 365–382. doi: 10.1044/1092-4388(2009/08-0205)

Dunn, J., and Kendrick, C. (1982). The speech of two- and three-year-olds to infant siblings: 'baby talk' and the context of communication. *J. Child Lang.* 9, 579–595. doi: 10.1017/S030500090000492X

Dunn, L. M., Theriault-Whalen, C. M., and Dunn, L. M. (1993). *Échelle de Vocabulaire en Images Peabody. Adaptation Française du Peabody Picture Vocabulary Test-Revised*. Toronto, ON: PSYCAN.

Engelhardt, P. E., Corley, M., Nigg, J., and Ferreira, F. (2010). The role of inhibition in the production of disfluencies. *Mem. Cognit.* 38, 617–628. doi: 10.3758/MC.38.5.617

Engelhardt, P. E., Nigg, J. T., and Ferreira, F. (2013). Is the fluency of language outputs related to individual differences in intelligence and executive function? *Acta Psychol. (Amst.)* 144, 424–432. doi: 10.1016/j.actpsy.2013.08.002

Ervin-Tripp, S. M. (1978). Some features of early child-adult dialogues. *Lang. Soc.* 7, 357–373. doi: 10.1017/S0047404500005777

Ervin-Tripp, S. M., Guo, J., and Lampert, M. (1990). Politeness and persuasion in children's control acts. *J. Pragmat.* 14, 307–331. doi: 10.1016/0378-2166(90)90085-R

Espy, K. A., and Cwik, M. F. (2004). The development of a trial making test in young children: the TRAILS-P. *Clin. Neuropsychol.* 18, 411–422. doi: 10.1080/138540409052416

Fagan, J. F., Holland, C. R., and Wheeler, K. (2007). The prediction, from infancy, of adult IQ and achievement. *Intelligence* 35, 225–231. doi: 10.1016/j.intell.2006.07.007

Favre, C. P., and Maeder, P. (2002). *Actes de Langage et Reconnaissance D'intention: la Logique Interlocutoire Appliquée à Des Conversations Avec Une Patiente Souffrant de Lésions Bifrontales Traumatiques*. Paris: Editions L'Harmattan.

Filippova, E., and Astington, J. W. (2008). Further development in social reasoning revealed in discourse irony understanding. *Child Dev.* 79, 126–138. doi: 10.1111/j.1467-8624.2007.01115.x

Gathercole, S. E. (2000). Limitations in working memory: implications for language development. *Int. J. Lang. Commun. Disord.* 35, 95–116. doi: 10.1080/136828200247278

Gathercole, S. E., and Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *J. Mem. Lang.* 29, 336–360. doi: 10.1016/0749-596X(90)90004-J

Gathercole, S. E., Tiffany, C., Briscoe, J., and Thorn, A. (2005). Developmental consequences of poor phonological short-term memory function in childhood: a longitudinal study. *J. Child Psychol. Psychiatry* 46, 598–611. doi: 10.1111/j.1469-7610.2004.00379.x

Grice, H. P. (1975). "Logic and conversation," in *Studies in Syntax, Vol. 3: Speech Acts*, eds P. Cole and J. L. Morgan (New York, NY: Academic Press), 41–58.

Hongwanishkul, D., Happaney, K. R., Lee, W. S., and Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Dev. Neuropsychol.* 28, 617–644. doi: 10.1207/s15326942dn2802_4

Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Dev. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Humphries, T., Koltun, H., Malone, M., and Roberts, W. (1994). Teacher-identified oral language difficulties among boys with attentional problems. *J. Dev. Behav. Pediatr.* 15, 92–98. doi: 10.1097/00004703-199404000-00005

Jagot, L., Marlier, N., Tissier, A. C., Patin, V., Azouvi, P., Le Mestric, L., et al. (2001). "Discours conversationnel et procédural chez le sujet traumatisé crânien sévère: étude conjointe de deux outils d'analyse clinique," in *Neuropsychologie et pragmatique*, ed C. Peter-Favre (Paris: L'Harmattan), 75–108.

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99, 122–149. doi: 10.1037/0033-295X.99.1.122

Landau, S., and Milich, R. (1988). Social communication patterns of attention-deficit-disordered boys. *J. Abnorm. Child Psychol.* 16, 69–81. doi: 10.1007/BF00910501

Lehto, J. E., Juujärvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Lemelin, J.-P., and Boivin, M. (2007). *Mieux Réussir dès la Première Année: L'importance de la Préparation à L'école*. Québec: Institut de la Statistique Québec.

Levy, B. J., and Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends Cogn. Sci.* 6, 299–305. doi: 10.1016/S1364-6613(02)01923-X

Lezak, M. D., Howieson, D. B., Bigler, E. D., and Tranel, D. (2012). *Neuropsychological Assessment. 5th Edn.* New York, NY: Oxford University Press.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., and Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Dev. Sci.* 7, 297–307. doi: 10.1111/j.1467-7687.2004.00349.x

Martin, I., and McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain Lang.* 85, 451–466. doi: 10.1016/S0093-934X(03)00070-1

Martinez, M. A. (1987). Dialogues among children and between children and their mothers. *Child Dev.* 58, 1035–1043. doi: 10.2307/1130544

McDonald, S. (2000). Exploring the cognitive basis of right-hemisphere pragmatic language disorders. *Brain Lang.* 75, 82–107. doi: 10.1006/brln.2000.2342

McDonald, S., and Pearce, S. (1996). Clinical insights into pragmatic theory: frontal lobe deficits and sarcasm. *Brain Lang.* 53, 81–104. doi: 10.1006/brln.1996.0038

McDonald, S., and Pearce, S. (1998). Requests that overcome listener reluctance: impairment associated with executive dysfunction in brain injury. *Brain Lang.* 61, 88–104. doi: 10.1006/brln.1997.1846

McKown, C. (2007). Concurrent validity and clinical usefulness of several individually administered tests of children's social–emotional cognition. *J. Clin. Child Adoles. Psychol.* 36, 29–41. doi: 10.1080/15374410709336566

Meng, X.-L., Rosenthal, R., and Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychol. Bull.* 111, 172–175. doi: 10.1037/0033-2909.111.1.172

Monette, S., and Bigras, M. (2008). La mesure des fonctions exécutives chez les enfants d'âge préscolaire. *Can. Psychol.* 49, 323–341. doi: 10.1037/a0014000

Nilsen, E. S., and Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cogn. Psychol.* 58, 220–249. doi: 10.1016/j.cogpsych.2008.07.002

Nippold, M. A. (1985). Comprehension of figurative language in youth. *Top. Lang. Disord.* 5, 1–20. doi: 10.1097/00011363-198506000-00004

Norbury, C. F., Nash, M., Baird, G., and Bishop, D. V. M. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the Children's communication checklist-2. *Int. J. Lang. Commun. Disord.* 39, 345–364. doi: 10.1080/13682820410001654883

Ostrov, J. M., Pilat, M. M., and Crick, N. R. (2006). Assertion strategies and aggression during early childhood: a short-term longitudinal study. *Early Child. Res. Q.* 21, 403–416. doi: 10.1016/j.ecresq.2006.10.001

Owens, R. E. (2011). *Language Development: An Introduction. 8th Edn*. Upper Saddle River, NJ: Pearson Education.

Ozonoff, S. (2001). "Advances in the cognitive neuroscience of autism," in *Handbook of Developmental Cognitive Neuroscience*, eds C. A. Nelson and M. Luciana (Cambridge, MA: MIT Press), 537–548.

Patterson, G. R., Littman, R. A., and Bricker, W. (1967). Assertive behavior in children: A step toward a theory of aggression. *Monogr. Soc. Res. Child Dev.* 32, 1–43. doi: 10.2307/1165737

Pellegrini, A. D., Brody, G. H., and Stoneman, Z. (1987). Children's conversational competence with their parents. *Dis. Process.* 10, 93–106. doi: 10.1080/01638538709544661

Raaijmakers, M. A. J., Smidts, D. P., Sergeant, J. A., Maassen, G. H., Posthumus, J. A., Van Engeland, H., et al. (2008). Executive functions in preschool children with aggressive behavior: Impairments in inhibitory control. *J. Abnorm. Child Psychol.* 36, 1097–1107. doi: 10.1007/s10802-008-9235-7

Rasmussen, C., Talwar, V., Loomes, C., and Andrew, G. (2008). Brief report: Lie-telling in children with fetal alcohol spectrum disorder. *J. Pediatr. Psychol.* 33, 220–226. doi: 10.1093/jpepsy/jsm069

Reisinger, L. M. (2011). *Exploring the Behavioral and Cognitive Profile of Children with Autism and Children with Pragmatic Language Impairment*. Doctoral Dissertations, McGill University.

Russell, R. L., and Grizzle, K. L. (2008). Assessing child and adolescent pragmatic language competencies: toward evidence-based assessments. *Clin. Child Fam. Psychol. Rev.* 11, 59–73. doi: 10.1007/s10567-008-0032-1

Sachs, J., Donnelly, J., Smith, C., and Bookbinder, J. D. (1991). Preschool children's conversation intrusions: behavior and metapragmatic knowledge. *Dis. Process.* 14, 357–372. doi: 10.1080/01638539109544790

Sattler, J. M. (2008). *Assessment of Children: Cognitive Foundations 5th Edn.* San Diego, CA: Jerome M. Sattler Publisher.

Schuh, J. M. (2012). *Pragmatic Language Abilities: Working Memory Influences on Mutual Information.* Doctoral Dissertations, University of Connecticut.

Smith, P. K. (2003). "Play and peer relations," in *An Introduction to Developmental Psychology*, eds A. Slater and G. Bremner (Malden, MA: Blackwell Publishing), 311–333.

Spector, C. C. (1996). Children's Comprehension of Idioms in the Context of Humor. *Lang. Speech Hear. Serv. Sch.* 27, 307.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence, 3rd Edn.* San Antonio, TX: Psychological Corporation.

Welsh, M. C., Pennington, B. F., and Groisser, D. B. (1991). A normative-developmental study of executive function: a window on prefrontal function in children. *Dev. Neuropsychol.* 7, 131–149. doi: 10.1080/87565649109540483

Zelazo, P. D. (2006). The dimensional change card sort (DCCS). *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Zelazo, P. D., and Müller, U. (2005). "Hot and cool aspects of executive function: relations in early development," in *Young Children's Cognitive Development: Interrelationships Among Executive Functioning, Working Memory, Verbal Ability, and Theory of Mind*, eds W. Schneider, R. Schumann-Hengsteler, and B. Sodian (Mahwah, NJ: Erlbaum), 71–93.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 January 2014; accepted: 04 March 2014; published online: 20 March 2014.
Citation: Blain-Brière B, Bouchard C and Bigras N (2014) The role of executive functions in the pragmatic skills of children age 4–5. Front. Psychol. ***5****:240. doi: 10.3389/fpsyg.2014.00240*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Blain-Brière, Bouchard and Bigras. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 20 May 2014
doi: 10.3389/fpsyg.2014.00469

The union of narrative and executive function: different but complementary

Margaret Friend* and Raven Phoenix Bates

Department of Psychology, San Diego State University, San Diego, CA, USA

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Nobuyuki Jincho, RIKEN Brain Science Institute, Japan
Jeffrey Coldren, Youngstown State University, USA

****Correspondence:***
Margaret Friend, Department of Psychology, San Diego State University, 6505 Alvarado Road, Ste. 101, San Diego, CA 92120, USA
e-mail: mfriend@mail.sdsu.edu

Oral narrative production develops dramatically from 3 to 5 years of age, and is a key factor in a child's ability to communicate about the world. Concomitant with this are developments in executive function (EF). For example, executive attention and behavioral inhibition show marked development beginning around 4 years of age. Both EF and oral narrative abilities have important implications for academic success, but the relationship between them is not well understood. The present paper utilizes a cross-lagged design to assess convergent and predictive relations between EF and narrative ability. As a collateral measure, we collected a Language Sample during 10 min of free play. Language Sample did not share significant variance with Narrative Production, thus general language growth from Wave 1 to Wave 2 cannot account for the predictive relations between EF and Narrative. Our findings suggest that although EF and Narrative ability appear independent at each Wave, they nevertheless support each other over developmental time. Specifically, the ability to maintain focus at 4 years supports subsequent narrative ability and narrative ability at 4 years supports subsequent facility and speed in learning and implementing new rules.

Keywords: executive function, narrative, attention, inhibition, preschool children

INTRODUCTION

Storytelling is integral to human culture: the ability to express a story using pictures and relate it to life is the essence of creating shared meaning. Oral narrative production develops dramatically from 3 to 5 years of age. Concomitant with this development are developments in executive function. For example, executive attention and behavioral inhibition show marked development beginning around 4 years of age. Both executive function and oral narrative abilities have important implications for academic success, but the relationship between them is not well understood.

One form of oral narrative is emergent reading, which occurs when children tell a story using a picture book for support (Sulzby, 1985; Valencia and Sulzby, 1991). Curenton and Justice (2004) found significant increases in the use of conjunctions and verbs in the narratives of preschoolers from 3 to 5 years of age. Story grammar also undergoes maturation during this.

Although children as young as 3 can between past and present tense, they rarely use past tense when telling a story. Tense marking improves along with the use of verbs and conjunctions by age 5 and this contributes to the ability to convey action and organize events in a coherent sequence (Berman and Slobin, 1994). In addition, Nicolopoulou and Richner (2007) found that at age 3 children often focus on physical aspects of characters whereas at age 4, character descriptions include some goal-related action and by age 5, children express a more complex representation of characters in their story telling (Nicolopoulou and Richner, 2007).

Narratives are a product of increasing linguistic sophistication over the preschool period (Kaderavek and Sulzby, 2000) and there is a complex relation between early narratives, language proficiency, and theory of mind (ToM). In a classic paper, Astington and Jenkins (1999) showed that the relation between language and ToM is unidirectional: early language predicts later ToM but early ToM does not predict later language. Charman and Shmueli-Goetz (1998) confirmed a strong relation between language and ToM but found a more limited relation between ToM and narrative: ToM was associated with referential strategy but not with mental state terms, length, complexity, or story structure. Recent work supports this circumscribed view of the relation between ToM and narrative. For example, Fernández (2013) found that ToM explained a small but significant portion of the variance in pragmatic language in children's narratives beyond variance explained by gender and language proficiency. Similarly Ketelaars et al. (2012) found that false belief understanding explained 7% of the variance in the narrative productivity (number of grammatical units, clauses, and MLU) beyond variance explained by language but did not account for variance in story organization or cohesion. This emphasizes the importance of selecting an approach to coding that focuses on the aspects of narrative under investigation. In the present research, we are particularly interested in aspects of narrative production that are likely to be associated with executive function.

Cobo-Lewis et al. (2002) developed a narrative complexity scale to assess narrative construction across languages in bilingual acquisition. This scale captures several aspects of narrative structure that are particularly likely to support and be supported by the development of executive function: memory for story elements, sequencing, demarcating the story with a clear beginning, middle, and end, and using complex syntax. This scale distinguished monolinguals from bilinguals on linguistic elements but

not on memory, sequencing, and structure suggesting that these components of narration are not confounded with language proficiency in typically developing children. Using a similar approach focusing on thematic aspects of children's narratives, Ilgaz and Aksu-Koç (2005) found clear improvement in structure from 3 to 5 years of age.

A review of the literature by Mar (2004) found evidence for a network of frontal, temporal, and cingulate areas supporting story comprehension and production. Narrative production and comprehension require substantial organizational skill and are particularly dependent on frontal cortical activation. Troiani et al. (2008) found support for this thesis in a magnetic imaging study of young adults narrating the children's picture story, "Frog, Where Are You" (Mayer, 1969). Peak activations were obtained bilaterally in the inferior frontal cortex as well as the temporal-parietal region and visual association cortex. Troiani et al. concluded that the bilateral frontal activation reflected the top-down organization that is necessary to construct an extended narrative. However the results also suggest a larger network that supports memory for story components, inferential meaning, and story organization.

Concomitant with the emergence of narrative ability, goal-directed action improves dramatically. The psychological processes underlying goal-directed action are referred to collectively as executive function (Zelazo et al., 2003) and there is consensus that substantial changes in executive function occur between 3 and 6 years of age (Carlson, 2003, 2005; Zelazo et al., 2003; Bunge and Zelazo, 2006; Crone et al., 2006; Garon et al., 2008; Moriguchi and Hiraki, 2009; Diamond, 2013). Executive attention, behavioral inhibition, and working memory are foundational higher-level processes that develop in early childhood (Best and Miller, 2010) although other recent work characterizes this triumvirate as set shifting (the ability to shift between rule sets), inhibition, and working memory (Miyake et al., 2000; Garon et al., 2008).

There is substantial theoretical overlap between these processes and shared variance in the tasks that tap them (Stelzer et al., 2014). For example, Best and Miller (2010) place the well-researched Dimensional Change Card Sort (DCCS) task squarely in the domain of complex behavioral inhibition whereas Garon et al. (2008) classify it as a set-shifting task. Further, inhibition tasks often place demands on working memory such that inhibition and working memory are not fully dissociable. Similarly, a recent study of the factor structure of executive function suggests that, in early childhood, set shifting and inhibition are not fully dissociable processes (Van der Ven et al., 2013). This, according to Miyake et al. (2000), is the problem of task impurity: each executive process operates on other processes. Nevertheless, these processes show compelling developmental change in the preschool period and have implications for subsequent achievement.

For the purposes of the present paper, we briefly review findings on the age-related change observed in this period with a focus on executive attention and inhibition. For example, Jones et al. (2003) found evidence of improvements in behavioral inhibition between 3 and 4 years of age on the Simple Simon Task. Children were instructed to follow the command of one large toy animal but not another. Error rates decreased between 3 and 4 years of age and, at age 4, children's response times incremented after making an error whereas this marker of error recognition was not evident in younger children.

On the DCCS task, Zelazo et al. (2003) taught children two sets of rules for sorting a set of cards: one based on shape and one based on color. They found that 3-year-olds understood each set of rules but failed to switch between them. Instead, the first set of rules learned determined the prepotent response on the task. Zelazo et al. (Zelazo and Frye, 1998; Zelazo et al., 2003) interpret this finding as evidence of a failure to reflect on the rules in relation to one another. Other accounts focus on conceptual redescription (Perner and Lang, 2002), latent vs. active memory (Munakata, 2001, 2004), and a failure to disengage attention from a previous rule set (Kirkham and Diamond, 2003). In sum, one can understand the difficulty of 3-year-olds on the DCCS and similar tasks as a problem of thinking about something in two ways simultaneously or, complementarily, as a difficulty of selective attention (Garon et al., 2008). A general finding is that the youngest children perseverate on the first rule pair to which they are exposed. Four- and five-year-olds, in contrast, are significantly more able to resist the prepotent response to the first rule (Zelazo and Jacques, 1996).

In Luria's Tapping Task, children are instructed to tap twice if the experimenter taps once and to tap once if the experimenter taps twice. Like the DCCS, this task requires that children keep both rules in mind simultaneously. In addition, it requires that children inhibit the prepotent tendency to imitate the experimenter. Accuracy on the task improves from 3.5 to 7 years of age (Diamond and Taylor, 1996). Recently, Clark et al. (2013) charted the trajectory of response inhibition and set shifting from 3 to 5 years of age. There was a clear improvement in accuracy on both measures and a reduction in response times.

A different approach, designed to capture individual differences in attention, the Child Attention Network Task (ANT; Rueda et al., 2004), was developed as an extension of the adult flanker task. Colorful fish appear on a screen and the child must "feed" the central fish using the arrow keys on the keyboard. To succeed, the child must focus on the direction that the fish is facing and, in incongruent trials, resist responding based on the orientation of the many other fish (flankers). Reaction time and accuracy improves with age across trial types (congruent, incongruent) and is significantly poorer for incongruent trials. Taken together, the results from these tasks indicate that executive function improves markedly during the period from age 3 to 5 with both qualitative and quantitative change apparent between 4 and 5 years of age.

Executive function, like narrative production, is associated with ToM (Perner and Lang, 1999) and inhibition and working memory are central to this relation (Carlson et al., 2002). Thus, speculatively, relations between executive function and narrative are likely to share variance with ToM through the domain general mechanisms of inhibition and working memory. Also like narrative production, the development of executive function has been associated with development in frontal cortical function (Perner and Lang, 1999). Additionally, improvements in executive function correlate with myelination and branching in the frontal lobe from infancy into middle childhood (Diamond and Taylor, 1996). However, executive function also depends upon a neural network

that extends across brain regions. Imaging studies suggest a network that is involved in the resolution of conflict (e.g., between a prepotent and appropriate response) comprised of the anterior cingulate and lateral prefrontal cortex (Fan et al., 2003, 2005) and the inferior frontal and parietal regions (Smith et al., 2004).

Performance on executive function tasks correlates with academic success in mathematics, reading, and writing. Clark et al. (2010) found that children who performed below average on measures of executive planning, attention, and inhibition at age 4 also performed below average on math skills at the first grade level. Interestingly, set shifting did not correlate with any other measure of executive function or with math achievement. Nevertheless, it is clear that set shifting is a central component of executive function. Indeed there has been substantial recent work indicating that set shifting may be an important component of dual language acquisition, supporting the ability to transition between languages and moreover, that dual language acquisition supports precocious development in set shifting (see Kroll et al., 2012, for a review). Although the fields of emergent literacy and executive control receive significant attention individually, relatively little research in typically developing children connects the two fields.

One possibility is that the effects of executive processes may be specific, supporting particular aspects of cognition at particular points in developmental time. Apropos of this hypothesis, Schneider et al. (2006) found that language and working memory at 36 months accounted for significant variance in executive control at 42 and 48 months suggesting that, in early childhood, both factors support subsequent developments in executive function, at least in the short term. However, consistent with the specificity hypothesis, planning, attention, and inhibition did not correlate with working memory and the strength of the prediction from early language to executive control decreased over time. Nevertheless, children with language deficits score significantly more poorly on both verbal and non-verbal executive function tasks than peers without language deficits (Bialystok and Feng, 2009) suggesting that typical language may be important to the development of executive function or, conversely, that typical executive function may be necessary to support language acquisition.

In a large longitudinal study of typically developing children, the National Institute of Child Health and Human Development (2003) found that sustained attention and behavioral inhibition at 54 months partially mediated the relation between home environment and cognitive, school readiness, language, and social outcomes. Other recent work suggests a direct relation between performance on the DCCS and later language and emergent literacy skills such as phonological sensitivity and print awareness (Bierman et al., 2008). In contrast, Coldren (2013) found that whereas DCCS scores correlated with math and district kindergarten exit scores, they did not account for significant variance in reading scores above that accounted for by age and school readiness.

These findings are consistent with the view that the executive processes underlying goal- directed behavior exhibit specificity of prediction: executive processes are not homogeneous but exhibit specific convergent and predictive relations that vary with developmental time. This view is consistent with Garon et al.'s (2008) model integrating unitary and componential approaches to executive function from a developmental perspective (see also Lehto et al., 2003; Huizinga et al., 2006 for alternate integrative models).

A handful of studies have examined the relation between executive processes and narrative production in brain-injured adults. Coelho et al. (1995) found that, in adults with traumatic brain injury (TBI), there was a significant correlation between story structure and executive function such that adults who produced incomplete episodes within the story were also less adept at learning the sorting rule in the Wisconsin Card Sorting Task. Additionally, TBI adults scored significantly lower than average on overall narrative cohesiveness.

In another study, Coehlo (2002) found that individuals with closed head injuries (CHI) produced less coherent episodes and used fewer words overall than adults without head injury and that narrative production was correlated with scores on the Wisconsin Card Sorting Test. Subsequently, Mozeiko et al. (2011) found no differences between a group of adults with TBI and a comparison group on measures of set shifting and inhibition. However, there were significant group differences in narrative organization such that the TBI group's narratives contained fewer content episodes. Further, in the TBI, but not the control, group the correlation between set shifting and story structure was significant. Thus narrative deficits and executive function deficits share variance in adults with closed head as well as traumatic brain injuries. This suggests that the two abilities depend on a shared underlying neural substrate and thus, it is reasonable to expect that executive processes and narrative ability are dependent over developmental time.

Of particular interest in the present research are the convergent and predictive relations between executive processes and narrative production from 4 to 5 years of age. Consider that, in order to tell a good story, a child must engage executive processes. She must maintain the rough structure of the story (what came before and what comes next and how these are related), concentrate on the complete telling of one segment at a time, and nimbly shift between one segment and the next in order to produce a well-structured narrative. In fact, story structure is what makes narrative cohere in a way that facilitates comprehension in a listener (Hudson and Shapiro, 1991; Shapiro and Hudson, 1991). Thus children must organize information in narratives into a set of causal chains that emphasize the temporal sequence and causal relevance of events within the story.

This, in conjunction with the fact that developments in narrative production emerge in concert with developments in executive function, suggests a potential developmental relation between executive processes and the ability to construct narratives. Further, evidence from imaging studies and from brain-injured adults suggests that the neurological networks that support executive function and narrative production are at least partially overlapping and that development in both domains is dependent upon the attention system. What is less clear is the direction of this relation over developmental time. Do executive processes emerge and mature in advance of proficient storytelling or does practice telling stories support the development of

executive processes? In a recent review, Diamond (2013) proposed an interdependent model of the relation between active, volitional inhibition and working memory. Successful inhibition requires the contribution of working memory. Similarly, and perhaps not as obviously, working memory requires inhibitory control: focusing the mind and remembering is dependent upon resistance to distraction. Of interest then is the nature of the developmental relation between executive function and narrative development.

The present research focuses on attention, inhibition, and narrative development in early childhood. Because we are interested in the convergent and predictive relations between narrative and executive function, we examine the period between 4 and 5 years of age retesting each participant within a 6 months window to observe how narrative supports executive function and how executive function supports narrative. Although it is possible to assess both executive function and narrative even earlier, we examine this period to minimize floor effects. It is expected that executive processes will correlate within each Wave. Further, we anticipate that each measure (attention, inhibition, and narrative) will correlate across Waves. Of particular interest are the correlations between executive processes and narrative production from Wave 1 to 2.

METHODS

PARTICIPANTS

A sample of 52 children between the ages of 48 and 60 months ($M = 53;27$) and their primary caregivers participated in the first Wave of this study. Ten children were excluded due to technical difficulties with the audio recorder (7) and general fussiness (3), leaving us with a final sample of 42 children ranging in income from $15,000 to $100,000 per year and with maternal education from 10 to 18 years. All participants were monolingual speakers of American English however roughly one-half reported exposure to a second language, reflecting our presence in a border region. A summary of sample demographics is presented in **Table 1**. A subset of the sample (38 caregiver-child dyads) returned to participate in the second Wave of the experiment when children were between the ages of 54 and 66 months ($M = 60;18$; see **Table 2**). Consistent with our primary objective of exploring how skills support one another over developmental time, performance was assessed across Waves using each child as his own control. Because the narrow interval between Waves resulted in some overlap in age (see **Figure 1** for a distribution of ages at each Wave), we assess performance on each dependent measure across Waves to insure inter-test interval is developmentally appropriate before proceeding to the cross-lagged analyses.

Table 1 | Distribution of selected demographic characteristics of participants Wave 1.

Characteristic	Boys ($N = 16$)	Girls ($N = 26$)	Total ($N = 42$)
MATERNAL EDUCATION			
High school or less	–	2 (7.7)	2 (4.8)
some college	4 (25.0)	7 (26.9)	11 (26.2)
college graduate	4 (25.0)	5 (19.2)	9 (21.4)
Post-baccalaureate	8 (50.0)	12 (46.1)	20 (47.6)
APPROXIMATE INCOME			
15,000–24,999	–	3 (11.5)	3 (7.1)
25,000–49,999	5 (31.2)	4 (15.4)	9 (21.4)
50,000–74,999	1 (6.2)	5 (19.2)	6 (14.3)
75,000–99,999	5 (31.2)	2 (7.7)	7 (16.7)
100,000–150,000	5 (31.2)	12 (46.1)	17 (40.5)
150,000+	–	–	–
MATERNAL ETHNICITY			
Asian	2 (12.5)	–	2 (4.8)
Black/not hispanic	1 (6.2)	2 (7.7)	3 (7.2)
Hispanic	3 (18.7)	9 (34.6)	12 (28.5)
White/not hispanic	10 (62.5)	15 (57.7)	25 (59.5)
mixed race	–	–	–
SECOND LANGUAGE			
No	9 (56.2)	11 (42.3)	20 (47.6)
Yes	7 (43.7)	15 (57.7)	22 (52.4)

Table 2 | Distribution of selected demographic characteristics of participants Wave 2.

Characteristic	Boys ($N = 14$)	Girls ($N = 24$)	Total ($N = 38$)
MATERNAL EDUCATION			
High school or less	–	2 (8.3)	2 (5.3)
Some college	3 (21.4)	6 (25.0)	9 (23.7)
College graduate	4 (28.6)	4 (16.7)	8 (21.0)
Post-baccalaureate	7 (50.0)	12 (50.0)	19 (50.0)
APPROXIMATE INCOME			
15,000–24,999	–	1 (4.2)	1 (2.6)
25,000–49,999	5 (35.7)	4 (16.7)	9 (23.7)
50,000–74,999	–	5 (20.8)	5 (13.2)
75,000–99,999	5 (35.7)	2 (8.3)	7 (18.4)
100,000–150,000	4 (28.6)	12 (50.0)	16 (42.1)
150,000+	–	–	–
MATERNAL ETHNICITY			
Asian	1 (7.1)	–	1 (2.6)
Black/not hispanic	1 (7.1)	2 (8.3)	3 (7.9)
Hispanic	3 (21.4)	8 (33.3)	11 (29.0)
White/not hispanic	9 (64.4)	14 (58.4)	23 (60.5)
Mixed race	–	–	–
SECOND LANGUAGE			
No	8 (57.1)	9 (37.5)	17 (44.7)
Yes	6 (42.9)	15 (62.5)	21 (55.3)

MEASURES

Narrative elicitation task

"Frog, where are you?" by Mercer Mayer (1969), a 24-page wordless picture book was used to elicit children's narratives. In the story a boy loses his frog and goes on a search to find him. Each page has a single picture of a scene in the story. The book has been used extensively to explore linguistic characteristics of narrative production in children and adults (Berman and Slobin, 1994).

A narrative complexity scale based on Cobo-Lewis et al. (2002) was used to code children's narratives. Previous research indicates that this approach captures aspects of narration that are not confounded with language proficiency in typically developing children. The scale included four subscales, summed to create a Narrative total score. The subscales were: (1) elements (e.g., a story that includes a loss, search, and discovery); (2) sequence

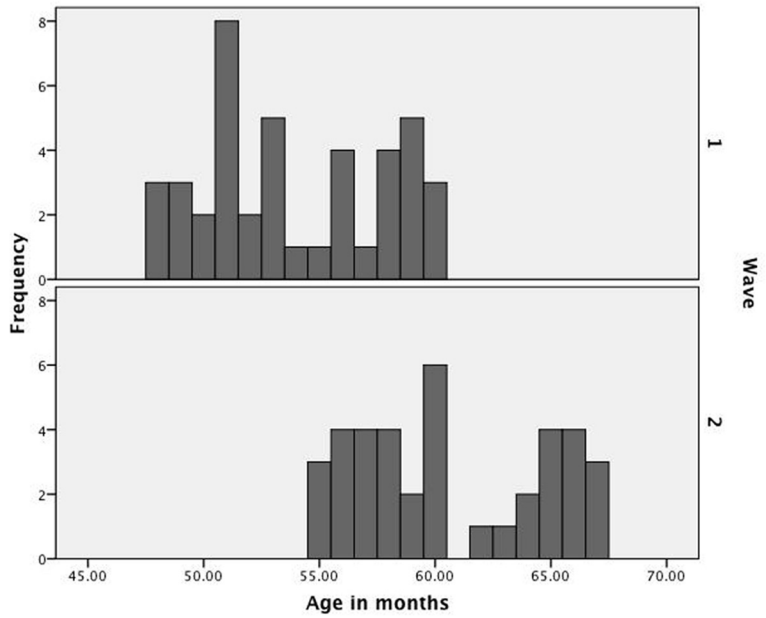


FIGURE 1 | Distribution of age across Waves 1 and 2.

(events organized in a causal sequence); (3) syntax (use of verb phrases, conjunctions, and/or adjective clauses); and (4) lexicon (use of a set of words specific to the story). Each subscale was scored separately on a scale from 0 to 12. For the elements subscale, the 12 primary story elements were identified (e.g., the frog is lost, the boy looks for the frog, the boy is sad, etc.). For the elements subscale, one point was assigned for each story element in the narrative. For the sequence subscale, scores were based on the completeness of the causal chain of events. The syntax and lexicon scores reflect a simple count of the number of complex constructions and relevant lexical items in the narrative (up to a total of 12). Inter-rater reliability between the primary coder and the second author was calculated for all of the stories and was >0.81 for Wave 1 and >0.85 for Wave 2.

The child attention network test

The Attention Network Test (ANT) assesses the alerting, orienting, and conflict resolution functions of attention in adults and has been adapted for use with children from 4 to 10 years of age (Rueda et al., 2004; Zelazo et al., 2013) and provides a broad measure of the functioning of the attention system. The ANT requires limited verbal instruction consistent with our goal of minimizing potential confounds with language proficiency. A bright yellow fish or a row of five yellow fish appears on the blue screen. The child is asked to help "feed" the central fish by pressing the arrow key corresponding to its orientation. In the neutral condition, only one fish appears. In the congruent condition, five fish appear all facing the same direction. In the incongruent condition, the flanking fish face the opposite direction of the central fish. Prior to presentation of the fish in each trial, one of four cue conditions appears on the screen. In the no cue condition, a fixation cross appears on the screen. In the central cue condition, an asterisk appears where the fixation cross was originally. In the single spatial cue condition, an asterisk appears above or below the fixation cross depending on where the fish will appear. In the double cue condition, two asterisks appear above and below the fixation cross. Resolving the conflict between the target and flanker fish in the incongruent condition has been shown to delay reaction times and activate regions of the lateral prefrontal cortex (Fan et al., 2003). Performance is strongly correlated with both the Block Design subtest of the WPPSI-III, although it is also correlates with the PPVT-IV indicating shared variance with language (Zelazo et al., 2013). Finally, the ANT reliably captures individual differences in attention (Posner et al., 2007).

During the test, the child sat 50 cm from a Dell PC with screen resolution 1280 × 1024. The child placed one finger each on the left and right arrows of the keyboard and used these to indicate the direction of the fish. The test consisted of a practice block of 24 trials, followed by two experimental blocks of 48 trials each. The child received audio feedback through the speakers on the computer. After a correct attempt, the child heard a "Woohoo!" audio-feedback while bubbles flowed from the middle fish's mouth. An incorrect attempt yielded no animation or audio-feedback. Following completion of each block, the child received a sticker as a reward. Reaction times and accuracy were recorded for each of the trials. In the second Wave, a DEX computer equipped with Windows operating system was used for the Child Attention Network Task. For the purposes of the present study, data were collapsed across congruent and incongruent trials to produce two summary scores: ANT Accuracy and ANT Reaction Time (RT).

Luria's Tapping Task. This task has been used to measure response inhibition in children 3½–7 years of age. The child and experimenter sat 45 cm across from each other at a table.

The experimenter held a wooden dowel 30.5 cm in length and 1 cm in diameter. The dowel was passed between child and experimenter to ensure that the child did not tap out of turn. The experimenter instructed the child to tap twice when the experimenter tapped once and to tap once when the experimenter tapped twice. A practice trial was given to insure the child understood the rules. If the practice trial was successful, the child moved on to two sessions of 16 pseudorandom trials each. Response latency and proportion of correct responses were measured. Children's responses were videotaped and coded offline by two experimenters at an inter-rater reliability 0.99.

The two executive function (EF) tasks, the Child ANT and Luria's Tapping Task were chosen from an array of candidate EF tasks for three reasons. First, both measures have been shown to capture changes in EF over the preschool period. Second, together they broadly assess attention itself as well as the known difficulty that children have keeping two things in mind simultaneously and resisting prepotent responses. Finally, both tasks have limited verbal demands thus reducing the potential for confounding performance with language proficiency.

Language sample

The child and caregiver were asked to play as they would at home for 10 min with a set of Duplos provided by the experimenter. This session was audiotaped for later transcription. Child language was transcribed into utterance units by one primary transcriber and the second author who completed one-third of the transcripts in common at an inter-rate agreement of 0.80. The transcripts were analyzed using the Systematic Analysis of Language Transcripts software (Miller and Iglesias, 2008). Two summary variables were computed: Number of Unique Words (NW) and Mean Length of Utterance in morphemes (MLU). These measures provide estimates of vocabulary size and grammatical complexity.

PROCEDURE

This research was approved the Institutional Review Board that oversees the protection of human research participants at San Diego State University. Primary caregivers contacted the lab by phone or email in response to advertisements posted on community-based Internet resources and in local daycare centers. Participants were introduced to the researcher in a 10-min warm-up period in the playroom of the lab while the caregiver filled out a consent form and a demographic questionnaire. All caregivers provided informed consent. Following the warm-up period, participants were taken to an adjacent testing room in the laboratory to complete the Child Attention Network Task (ANT). This room was equipped with a Dell PC on which the ANT program was installed. The child was seated on a chair and used the arrow keys on the keyboard to indicate their responses. The experimenter sat next to the child to explain the task, and behind the child during testing. Following each set of trials, a sticker was given to reward the child. Caregivers observed quietly from across the room throughout EF testing.

Following the Child ANT, the experimenter directed the caregiver and child to a second testing room equipped with a one-way mirror, a Sony Digital Video Camera Recorder Model DCR-TRV 350 in an adjacent room positioned behind the mirror, and a high-quality Audio Technica AT898 Subminiature Cardioid Condenser Lavalier Microphone housed discreetly in a conduit between the two rooms. The microphone recorded onto a Sony TCD-D7 DAT recorder.

For the next 10 min, caregivers engaged in free play with the child with a set of Duplos blocks. Next, the experimenter showed the child the picture book, "Frog, where are you?" by Mercer Mayer (1969), and asked the child to tell her a story using the pictures in the book. Finally, the child and experimenter completed Luria's Tapping Task (Diamond and Taylor, 1996). Tasks were completed in the same order for each participant. It was reasoned that the EF tasks were the most demanding so free play and storytelling were used to break up these tasks to insure compliance and optimal performance. Children also completed a school readiness measure as part of a larger study.

RESULTS

For each EF task, accuracy and speed were assessed at each Wave. Results from Wave 1 and Wave 2 were analyzed separately and then cross-panel correlations were run to assess the predictive relation from executive function to narrative production and from narrative production to executive function. It was expected that accuracy and speed would correlate across the two EF tasks at each Wave and that narrative production and EF would correlate across Waves.

WAVE 1

Descriptive statistics for the narrative production scores are presented in **Table 3**. All subscales were normally distributed and the full range of scores was utilized. The inter-item reliability coefficient for narrative production ($\alpha = 0.88$) was high, indicating good internal consistency. Descriptive statistics for latency and accuracy on the EF measures are presented in **Table 4**. Total Narrative scores (skew = 0.129, $SE = 0.365$), ANT accuracy (skew = −0.715, $SE = 0.365$), ANT latency (skew = 0.039,

Table 3 | Means, standard deviations, and ranges for narrative production scores at Wave 1.

Scale	*M*	*SD*	Range
Elements	6.02	2.32	1.0–11.0
Sequence	5.57	3.08	0–12.0
Syntax	4.76	2.67	0–10.0
Lexicon	7.48	1.70	3.0–11.0
Narrative total	23.83	8.56	7.0–40.0

Table 4 | Means, standard deviations, and ranges for tapping task and child ANT at Wave 1.

Measure	*M*	*SD*	Range
Tapping accuracy (%)	79.28	17.12	41.0–100.0
Tapping latency (ms)	1817.62	591.10	1130.0–4260.0
ANT accuracy	75.92	14.86	41.1–99.2
ANT latency (ms)	1464.79	290.25	845.0–2159.25

$SE = 0.365$) and Tapping accuracy (skew $= -0.742$, $SE = 0.365$) were normally distributed but Tapping latency exhibited a positive skew (skew $= 2.06$, $SE = 0.365$). A square root transform was performed on tapping latency scores to normalize the data. Findings were the same for the transformed and untransformed scores therefore we report on the untransformed data. Z-scores were calculated for all dependent measures for the purpose of detecting outliers. Visual inspection of the data revealed no outliers for Narrative or for the EF accuracy measures. A criterion of 2.5 *SD* from the mean was employed based on for the latency measures. Two outliers were identified with reaction times outside this window: 1 participant on the tapping task and 1 on the ANT task.

To determine whether age, preschool experience, and language proficiency and exposure influenced performance on narrative and executive function tasks at Wave 1, a MANOVA was conducted with Age, Number of Years in Preschool, and NW and MLU from the Language Sample as covariates, Sex and Second Language Exposure (yes/no) as fixed effects, and ANT latency and accuracy, Tapping latency and accuracy, and Narrative as dependent measures. Power for the full model was high (0.956) and the model was significant, $F_{(5,\ 32)} = 5.165$, $p = 0.001$, partial $\eta^2 = 0.450$, with an effect of Age, $F_{(5,\ 32)} = 3.055$, $p = 0.023$, partial $\eta^2 = 0.330$, but no other predictors or covariates reached significance. The analysis repeated with outliers removed yielded no difference in findings. Thus, relations between Narrative and EF cannot be explained by variance due to language proficiency or exposure.

The fact that we did not find a significant relation between the Language Sample and Narrative suggests that our narrative coding system minimized any confound with language proficiency. Narrative storytelling differs from spontaneous language in that storytelling is constrained to a specific subset of lexical items and constructions. Further, in the present study, the Narrative score also reflects the ability to structure language in a causally relevant way that captures all of the salient elements of the story. Thus the total score captures not only words and constructions but also organization and memory. The absence of an effect of language exposure is also not surprising: all participants were monolingual speakers of American English despite some second language exposure. To adequately assess the effects of language exposure, a design including control and comparison groups based upon a fine-grained assessment of the sources and durations of exposure would be necessary.

We proceed with a consideration of the zero-order correlations between EF and Narrative measures as well as the partial correlations controlling for Age. The correlations for the EF tasks and Narrative are presented in **Table 5A**. As expected, accuracy on the two EF tasks was significantly and positively correlated however when controlling for Age, this relation was not significant. In addition, on the Tapping Task, accuracy was significantly and negatively correlated with latency for both zero-order and partial correlations.

A second set of zero-order and partial correlations was computed with outliers on the reaction time measures removed (see **Table 5B**). These correlations differed in several important ways from correlations based on the full data set. First, ANT accuracy and Tapping latency were moderately related. Second the expected relation emerged between the two latency measures.

Table 5A | Wave 1 Correlations (outliers included).

		ANT accuracy	ANT latency	Narrative	Tapping accuracy	Tapping latency
ANT accuracy	Zero-order	1	−0.003	**−0.328**	**0.674**	−0.253
	Partial		−0.005	0.010	0.094	0.144
ANT latency	Zero-order		1	0.000	−0.036	0.150
	Partial			−0.178	**0.524**	**−0.238**
Narrative	Zero-order			1	0.133	−0.064
	Partial				0.034	−0.041
Tapping accuracy	Zero-order				1	**−0.446**
	Partial					**−0.244**

Correlations significant at $p < 0.05$ bolded. $N = 42$.

Table 5B | Wave 1 correlations (outliers excluded).

		ANT accuracy	ANT latency	Narrative	Tapping accuracy	Tapping latency
ANT accuracy	Zero-order	1	−0.026	−0.009	**0.727**	**−0.420**
	Partial		0.082	−0.153	**0.644**	**−0.351**
ANT latency	Zero-order		1	−0.263	−0.046	**0.416**
	Partial			−0.220	0.072	**0.379**
Narrative	Zero-order			1	0.173	−0.307
	Partial				0.047	−0.246
Tapping accuracy	Zero-order				1	**−0.433**
	Partial					**−0.363**

Correlations significant at $p < 0.05$ bolded. $N = 40$.

Third, the negative relation between Narrative scores and ANT accuracy for the zero-order correlations did not replicate when outliers were removed. Finally, the pattern of correlation was consistent across zero-order and partial correlations. That is, controlling for age no longer altered the pattern of results and, as predicted, the accuracy and latency measures were correlated across the two EF tasks.

Contrary to expectations, there was no relation between accuracy and latency on the ANT. This perhaps points to differences in the way that the two EF tasks tap executive processes. In the Tapping Task, which involves both working memory to keep track of the rule and inhibitory control to resist imitating the experimenter, speed may be essential to accurate performance since a delay would place additional demands on working memory. The ANT, in contrast, primarily assesses executive attention: memory demands are limited and responding quickly is less important to performance than maintaining focus on the target.

WAVE 2

For Narrative, all subscales were normally distributed and utilized the full range of scores (see **Table 6**). The inter-item reliability

Table 6 | Means, standard deviations, and ranges for narrative production scores at Wave 2.

Scale	*M*	*SD*	Range
Elements	6.60	1.85	3.0–10.0
Sequence	7.03	2.90	3.0–12.0
Syntax	4.82	2.28	1.0–10.0
Lexicon	8.29	1.27	5.0–11.0
Narrative total	26.76	6.51	15.0–40.0

coefficient was high ($\alpha = 0.762$), demonstrating good internal consistency and inter-rater reliability was also high ($\alpha > 0.85$). Total Narrative scores were normally distributed (skew = 0.379, $SE = 0.393$) as were ANT accuracy (skew = −0.790, $SE = 0.393$) and latency (skew = 0.503, $SE = 0.393$). Tapping accuracy exhibited a negative skew whereas Tapping latency exhibited a positive skew (skew = −1.307, $SE = 0.393$, and skew = 1.063, $SE = 0.393$, respectively) and inter-rater reliability for the Tapping Task was high ($\alpha = 0.99$). A square transform was performed on tapping accuracy and a square root transform on tapping latency to normalize the data. Findings were identical for the transformed and untransformed scores therefore we report on the untransformed data. **Table 7** presents descriptive statistics for the latency and accuracy scores for the EF measures. As in Wave 1, Z- scores were calculated for all dependent measures for the purpose of detecting outliers using a criterion of 2.5 *SD* from the mean. One outlier was identified with a Tapping accuracy score outside this window. No outliers were identified on the other measures.

To determine whether age, preschool experience, and language proficiency and exposure influenced performance on narrative and executive function tasks at Wave 2, we conducted a MANOVA with Age at Wave 2, Number of Years in Preschool, and NW and MLU from the Language Sample as covariates, Sex and Second Language Exposure (yes/no) as fixed effects, and ANT latency and accuracy, Tapping latency and accuracy, and Narrative as dependent measures. Power for the full model was high (0.924) and, although the model was significant, $F_{(5,\ 19)} = 3.0, p = 0.037$, partial $\eta^2 = 0.591$ there was no effect of any covariate or predictor. The model with the outlier removed was not significant. As at Wave 1, relations between Narrative and EF cannot be explained by variance due to language proficiency or exposure. We now consider the zero-order correlations between EF and Narrative. Removal of the outlier did not alter the pattern of findings and results are reported on the full dataset in **Table 8**.

As expected, and consistent with Wave 1, accuracy on the two EF tasks was significantly and positively correlated. This, in conjunction with the absence of significant variance attributable to age, suggests that the two EF measures begin to converge in their assessment of executive processes by about 5 years of age. In contrast to Wave 1 however, there was no relation between latency and accuracy on either EF task at 5 years of age, although there was a significant relation between latency on the Tapping Task and accuracy on the ANT. Recall that the only significant relation between accuracy and latency was for the Tapping Task in Wave 1. There was a marginal correlation [$r_{(36)} = 0.297,\ p = 0.08$] between ANT accuracy and Narrative suggesting that the ability to focus attention may be related to narrative production. Of particular interest however, are the cross-lagged correlations from Wave 1 to Wave 2.

Table 7 | Means, standard deviations, and ranges on tapping task and ANT at Wave 2.

Measure	*M*	*SD*	Range
Tapping accuracy (%)	92.44	8.65	69.0–100.0
Tapping latency (ms)	1705.26	490.74	1200.0–3180.0
ANT accuracy	86.57	12.17	57.3–100.0
ANT latency (ms)	1296.32	250.56	798.9–1881.42

Table 8 | Wave 2 zero-order correlations.

	ANT accuracy	ANT latency	Narrative	Tapping accuracy	Tapping latency
ANT accuracy	1	−0.168	−0.026	**0.437**	**−0.459**
ANT latency		1	−0.297	0.191	0.290
Narrative			1	0.184	−0.203
Tapping accuracy				1	−0.141
Tapping latency					1

Correlations significant at $p < 0.05$ bolded. $N = 38$.

LONGITUDINAL ANALYSES

Before proceeding with the longitudinal analyses, it is important to note that there was 10% attrition from Wave 1 to Wave 2. Included in this attrition were two outliers on the Wave 1 measures. Consequently, these outliers were not part of the sample at Wave 2 and do not contribute data to the longitudinal analyses. Removal of the single outlier at Wave 2 did not alter the pattern of longitudinal findings. Therefore we report all longitudinal findings, including the cross-lagged correlations, on the full sample from Wave 2. We expected each of the measures at Wave 1 to correlate with the same measure at Wave 2. In general, this expectation was supported. Narrative production scores at Wave 1 marginally correlated with narrative production scores at Wave 2. For the EF measures, ANT latency at Wave 1 significantly correlated with ANT latency at Wave 2 and ANT accuracy at Wave 1 significantly correlated with ANT accuracy at Wave 2. Tapping latency at Wave 1 marginally correlated with latency at Wave 2 and Tapping accuracy at Wave 1 significantly correlated with accuracy at Wave 2. The general picture is one of consistency over time in both EF and Narrative.

Next we evaluated the change in performance from Wave 1 to Wave 2 in each dependent measure to determine whether the interval between Waves was sufficient to inform our understanding of development in EF and Narrative (see **Table 9**). The change in performance was significant for ANT accuracy and latency, Tapping accuracy, and Narrative and marginally significant for Tapping latency. Taken together, the pattern indicates developmental change in individual children in EF and Narrative across

a 6-months window in the fifth year. Of particular interest are the cross-lagged relations between EF and Narrative. These were computed with and without outliers and the pattern of findings was comparable. Findings are reported on the full dataset (see **Table 10**).

Narrative at Wave 1 emerged as a significant predictor of Tapping latency at Wave 2, $r_{(36)} = -0.379$, $p = 0.022$, suggesting that practice producing meaningful narratives may support the ability to shift nimbly between responses on a task that taps working memory and inhibition. To explore this finding further, we examined the correlation of each Narrative subscale at Wave 1 with Tapping latency at Wave 2. Both the elements subscale, $r_{(36)} = -0.374$, $p = 0.025$, and the sequence subscale, $r_{(36)} = -0.387$, $p = 0.02$, emerged as significant predictors of subsequent Tapping latency. Importantly both the elements and sequence subscales place demands on working memory and inhibition to recall all of the relevant story elements and to organize them in a meaningful causal sequence. Thus children who are relatively good at constructing a narrative at age 4.5 are likely to be able to shift between arbitrary rules at age 5. There were no other significant relations between Narrative at Wave 1 and EF measures at Wave 2.

Turning to look at the prediction from EF to Narrative, the only significant prediction was from ANT accuracy at Wave 1 to Narrative at Wave 2, $r_{(36)} = 0.337$, $p = 0.044$. The better children were able to focus on the target and resist distraction at Wave 1, the more mature their narratives at Wave 2. We examined the correlation of each Narrative subscale at Wave 2 with ANT accuracy at Wave 1but found no significant effects other than for the total score.

To further clarify the developmental relation between Narrative and EF, partial correlations were calculated from Wave 1 to Wave 2 controlling for the influence of performance at Wave 1 on Wave 2 scores. Narrative production at Wave 1 remained significantly correlated with Tapping latency at Wave 2 even after controlling for Tapping latency at Wave 1 [$r_{(35)} = -0.380$, $p = 0.020$]. In addition, ANT accuracy at Wave 1 remained significantly correlated with Narrative at Wave 2 after controlling for Narrative at Wave 1 [$r_{(35)} = 0.362$, $p = 0.028$].

These results support the notion of bidirectional support between EF and Narrative over developmental time. Focusing and resisting distraction on the ANT in the fourth year predicts the ability to construct a causally coherent narrative in the fifth year, and the ability to construct a narrative in the fourth year predicts the speed with which children can follow arbitrary rules in the fifth year.

One concern was the potential for practice effects from Wave 1 to Wave 2 on narrative elicitation of the frog story. To account for potential practice effects we examined the correlation in narrative production across Waves, controlling for the difference in spontaneous language MLU and NW. The correlation was nonsignificant, suggesting that narrative production was not subject to practice effects over the 6 month testing interval.

Table 9 | Change in performance from Wave 1 to Wave 2.

Measure	Mean difference (Wave 2–Wave 1)	*SD*	$t_{(35)}$	*p*
ANT accuracy	0.10	0.10	6.05	0.001
ANT latency	−170.19	271.49	−3.76	0.000
Tap accuracy	0.12	0.14	5.15	0.000
Tap latency	−2.48	6.89	−2.10	0.086
Narrative	3.28	9.26	2.12	0.041

With outliers removed, the difference in Tap Latency is significant (p = 0.04) and the difference in Narrative is marginal (p = 0.07).

DISCUSSION

The ability to construct a narrative and components of executive function (e.g., the ability to focus attention, resist distraction, and shift nimbly between arbitrary rules) develop rapidly in the preschool period. Further, these skills are dependent upon overlapping neural substrates, particularly frontal lobe function, and deficits across these skill sets are observed in adults with traumatic and closed head injuries. Lastly, both sets of skills have been implicated in success in the early school years. In spite of these interesting parallels, the relation between narrative and executive function skills during this period has received little attention. The seminal question here is whether these are independent skill sets that just happen to develop concomitantly or whether there is a developmental relation between them such that executive function supports the development of narrative storytelling and practice constructing complex, causally coherent narratives supports development in executive function.

One issue that arises in assessing the relation between executive function and narrative ability is that, although there are many reasons to expect that the two skill sets might be related, causality is difficult to establish. Further complicating this picture is the fact that development can be heterochronous with skills that are deeply conceptually related developing on different

Table 10 | Wave 1–Wave 2 cross-lagged correlations (outliers included).

	W2 ANT accuracy	W2 ANT latency	W2 narrative	W2 tapping accuracy	W2 tapping latency
W1 ANT accuracy	**0.747**	−0.012	**0.337**	**0.632**	−0.216
W1 ANT latency	0.012	**0.476**	−0.060	0.229	0.282*
W1 narrative	0.125	−0.012	−0.317*	0.010	**−0.379**
w1tapping accuracy	**0.474**	0.231	0.209	**0.470**	−0.156
W1 tapping latency	−0.216	−0.239	−0.046	−0.224	0.285*

*Correlations significant at p < 0.05 bolded. *indicates a marginal correlation at p < 0.10. N = 38.*

timescales. Even though narrative ability and executive function develop across the preschool period, it is not necessarily the case that they do so in lock step. Some aspects of each skill set may develop before others and the relation between skill sets may be such that there is specificity in predictive relations over developmental time. That is, there is no compelling reason to think that all EF measures should equally share variance with the development of storytelling or that relations between EF and narrative should be apparent at any single point in time. For these reasons, we did not necessarily expect to see a relation between executive function and narrative at any one point in time but did anticipate predictive relations in our longitudinal analyses. We begin with a brief review of the primary convergent findings with Waves and then turn to a discussion of our longitudinal findings.

Consistent with our expectations, accuracy on EF measures converged at each Wave. However, contrary to our expectations, the relation between the two measures with regard to speed was much weaker such that we observed a relation between speed and accuracy for the Tapping Task, but not the ANT, in Wave 1 but not in Wave 2. We speculated that speed might be more important in the Tapping Task owing to memory demands. However, it is also the case that, across Waves, Tapping latencies were longer and more variable than ANT latencies and this variability may have contributed to the observed relation between speed and accuracy in Wave 1. At Wave 2, we found no relation between speed and accuracy within EF measures but a significant relation between Tapping latency and ANT accuracy. This effect is somewhat puzzling. This, taken together with the fact that Tapping latency was particularly variable, argues for caution in interpreting this relation between EF measures. With regard to the relation between executive function and narrative, our findings argue against a convergent relation at either 4.5 or 5 years of age. However, predictive relations between the two skill sets emerged in the longitudinal analyses.

Our findings revealed that more advanced narratives at 4.5 years of age were indicative of faster performance on the Tapping Task at 5 years of age. Importantly, this relation was not reciprocal: Tapping latency at 4.5 years of age did not predict narrative ability at age 5. This absence of reciprocity in addition to the fact that there was no convergent relation between Narrative and Tapping at either Wave constrains our interpretation. For example, if it were the case that the two measures correlate due to a third variable such as a shared neural substrate or synchronous developmental timing, we would expect to see convergent relations at each Wave as well as reciprocity in prediction. More likely, given the current evidence, the ability to structure a meaningful, causally coherent narrative supports the subsequent development of speed in responding to arbitrary rules and inhibiting prepotent responses. This finding is similar to recent findings suggesting that bilingualism supports set shifting performance (Soveri et al., 2011). Bilinguals must choose between languages or, put another way, between rule systems and response sets, in every conversation. It is thought that practice shifting between rules and responses underlies a bilingual advantage in executive function, particularly in tasks that involve set shifting. Similarly, we found that the better children were at constructing narratives at 4.5 years, the more quickly they were able to respond to a set of arbitrary rules at 5 years of age. Further, there was suggestive evidence that this relation was driven by children's competence in remembering all of the relevant story elements and organizing them in a causally coherent manner. Thus, skill at keeping track of and organizing key elements in storytelling, like being able to nimbly and appropriately shift between languages, appears to support subsequent speed in responding to arbitrary rules.

We also found a significant positive relationship between children's accuracy on the ANT at Wave 1 and Narrative at Wave 2. Like the relation between Tapping and Narrative, this relation was not reciprocal: Narrative at Wave 1 did not predict ANT accuracy at Wave 2. This finding provides further support for the notion that there are specific relations between narrative production and executive function across developmental time and that these relations reveal the ways in which the two skill sets support one another. This finding suggests that the ability to focus attention and resist distraction at 4.5 years confers benefits in the ability to construct a complex and coherent narrative at 5 years of age. Focusing attention, on what comes first, what comes next, who the relevant players are, and how events are related is key to telling a good story. Similarly, resisting distraction by peripheral information helps a narrator maintain the causal thread that is essential to constructing a meaningful narrative.

Taken together, these findings reveal an asynchronous relationship between executive function and narrative production. Importantly, the nature of this relationship depends upon the specific skills in question and upon the developmental time at which the skills are assessed. We found no strong evidence for a convergence of narrative and executive function skills at either 4.5 or 5 years of age. Rather, specific executive function skills predicted later narrative ability and narrative ability predicted subsequent specific, non-reciprocal, executive function skills. Narrative is not a component of executive function nor is it exclusively an outcome of language development. In fact, we found no relation between spontaneous language and either narrative or executive function. It should be noted, of course, that the particular relations observed between language and narrative will be dependent on the aspects of narrative production that are the focus of the coding scheme. In the present study, we chose an approach that emphasized inclusion of relevant story elements and causal structure as well as more linguistic aspects of storytelling such as syntax and story lexicon. In sum, we found that narrative and executive function are comprised of a set of skills that appear to develop asynchronously during the preschool period and that support subsequent development across skill sets. This finding is consistent with previous research revealing interdependency between executive function and theory of mind (Perner and Lang, 1999; Carlson et al., 2002). However, the present findings extend this work by showing that developments in executive function *per se* do not necessarily precede developments in narrative ability. Rather, there is a true interdependency such that developments in one domain support subsequent developments in the other. This finding is consistent with the work discussed earlier showing a bilingual advantage in some executive function tasks. Further, it extends Perner et al. (2002) approach to the relation between theory of mind and executive function to include the development of narrative ability. Finally, this approach is

consistent with Diamond's (2013) interdependent model of the relation between inhibition and working memory and reveals how such an account can conceptually integrate the many aspects of language and cognition that develop rapidly over the preschool period.

It is important to note that we focused the present research on a short window of time late in the fourth year when we expected to see marked development in both narrative and executive function. Our findings are suggestive of intriguing causal connections between these two skill sets. It will be interesting in future research to assess development across the preschool period to clarify these relations. In addition, the present sample size precluded more complex latent variable analyses. Indeed, although the power for the full models in our omnibus tests was high, power for our predictors was not owing to the small sample. These findings require replication with larger samples and modeling approaches to offer more definitive evidence on the relation between narrative and executive function across developmental time.

ACKNOWLEDGMENTS

We gratefully acknowledge all of the families that participated in our research. Preparation of this manuscript was supported by award #HD068458 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessary represent the official views of the NIH.

REFERENCES

Astington, J. W., and Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Dev. Psychol.* 35, 1311. doi: 10.1037/0012-1649.35.5.1311

Berman, R., and Slobin, D. I. (1994). *Relating Events in Narrative.* Hillsdale, NJ: L. Erlbaum Associates.

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Bialystok, E., and Feng, X. (2009). Language proficiency and executive control in proactive interference: evidence from monolingual and bilingual children and adults. *Brain Lang.* 109, 93–100. doi: 10.1016/j.bandl.2008.09.001

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., and Domitrovich, C. E. (2008). Executive functions and school readiness intervention: impact, moderation, and mediation in the Head Start REDI program. *Dev. Psychopathol.* 20, 821–843. doi: 10.1017/S0954579408000394

Bunge, S. A., and Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Curr. Dir. Psychol. Sci.* 15, 118–121. doi: 10.1111/j.0963-7214.2006.00419.x

Carlson, S. M. (2003). Executive function in context: development, measurement, theory, and experience. *Monogr. Soc. Res. Child Dev.* 68, 138–151. doi: 10.1111/j.1540-5834.2003.06803012.x

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M., Moses, L. J., and Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant Child Dev.* 11, 73–92. doi: 10.1002/icd.298

Charman, T., and Shmueli-Goetz, Y. (1998). The relationship between theory of mind, language and narrative discourse: an experimental study. *Curr. Psychol. Cogn.* 17, 245–271.

Clark, C., Pritchard, V., and Woodward, L. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Dev. Psychol.* 46, 1176–1191. doi: 10.1037/a0019672

Clark, C. A., Sheffield, T. D., Chevalier, N., Nelson, J. M., Wiebe, S. A., and Espy, K. A. (2013). Charting early trajectories of executive control with the shape school. *Dev. Psychol.* 49, 1481. doi: 10.1037/a0030578

Cobo-Lewis, A., Eilers, R., Pearson, B. Z., and Umbel, V. (2002). "Interdependence of Spanish and English knowledge in language and literacy among bilingual children," in *Language and Literacy in Bilingual Children*, eds D. K. Oller and R. Eilers (Clevedon: Multilingual Matters), 118–134.

Coehlo, C. (2002). Story narratives of adults with closed head injury and non-brain-injured adults: influence of socioeconomic status, elicitiation task, and executive functioning. *J. Speech Lang. Hear. Res.* 45, 1232–1248. doi: 10.1044/1092-4388(2002/099)

Coelho, C., Liles, B., and Duffy, R. (1995). Impairments of discourse abilities and executive functions in traumatically brain-injured adults. *Brain Inj.* 9, 471–477. doi: 10.3109/02699059509008206

Coldren, J. T. (2013). Cognitive control predicts academic achievement in kindergarten children. *Mind Brain Educ.* 7, 40–48. doi: 10.1111/mbe.12006

Crone, E. A., Donohue, S. E., Honomichl, R., Wendelken, C., and Bunge, S. A. (2006). Brain regions mediating flexible rule use during development. *J. Neurosci.* 26, 11239–11247. doi: 10.1523/JNEUROSCI.2165-06.2006

Curenton, S., and Justice, L. (2004). African American and Caucasian preschoolers' use of decontextualized language: literate language features in oral narratives. *Lang. Speech Hear. Serv. Sch.* 35, 240–253. doi: 10.1044/0161-1461(2004/023)

Diamond, A. (2013). Executive functions. *Ann. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Diamond, A., and Taylor, C. (1996). Development of an aspect of executive control: development of the abilities to remember what I said and to "Do as I say, not as I do." *Dev. Psychobiol.* 29, 315–334.

Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., and Posner, M. I. (2003). Cognitive and brain consequences of conflict. *Neuroimage* 18, 42–57. doi: 10.1006/nimg.2002.1319

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., and Posner, M. I. (2005). The activation of attentional networks. *Neuroimage* 26, 471–479. doi: 10.1016/j.neuroimage.2005.02.004

Fernández, C. (2013). Mindful storytellers: emerging pragmatics and theory of mind development. *First Lang.* 33, 20–46. doi: 10.1177/0142723711422633

Garon, N., Bryson, S. E., and Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31. doi: 10.1037/0033-2909.134.1.31

Hudson, J. A., and Shapiro, L. R. (1991). "From knowing to telling: the development of children's scripts, stories, and personal narratives," in *Developing Narrative Structure*, eds A. McCabe and C. Peterson (Hillsdale, NJ: Lawrence Erlbaum Associates), 89–136.

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Ilgaz, H., and Aksu-Koç, A. (2005). Episodic development in preschool children's play-prompted and direct-elicited narratives. *Cogn. Dev.* 20, 526–544. doi: 10.1016/j.cogdev.2005.08.004

Isquith, P., Crawford, J., Espy, K., and Gioia, G. (2005). Assessment of executive function in preschool-aged children. *Ment. Retarrd. Dev. Disabil. Res. Rev.* 11, 209–215. doi: 10.1002/mrdd.20075

Jones, L., Rothbart, M., and Posner, M. (2003). Development of executive attention in preschool children. *Dev. Sci.* 6, 498. doi: 10.1111/1467-7687.00307

Kaderavek, J. N., and Sulzby, E. (2000). Narrative production by children with and without specific language impairment: oral narratives. *J. Speech Lang. Hear. Res.* 43, 34–49.

Ketelaars, M. P., Jansonius, K., Cuperus, J., and Verhoeven, L. (2012). Narrative competence and underlying mechanisms in children with pragmatic language impairment. *Appl. Psycholinguist.* 33, 281–303. doi: 10.1017/S014271641100035X

Kirkham, N. Z., and Diamond, A. (2003). Sorting between theories of perseveration: performance in conflict tasks requires memory, attention and inhibition. *Dev. Sci.* 6, 474–476. doi: 10.1111/1467-7687.00303

Kroll, J. F., Dussias, P. E., Bogulski, C. A., and Valdes Kroff, J. R. (2012). Juggling two languages in one mind: what bilinguals tell us about language processing and its consequences for cognition. *Psychol. Learn. Motiv. Adv. Res. Theory* 56, 229. doi: 10.1016/B978-0-12-394393-4.00007-8

Lehto, J. E., Juujärvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Mar, R. A. (2004). The neuropsychology of narrative: story comprehension, story production and their interrelation. *Neurophsychologia* 42, 1414–1434. doi: 10.1016/j.neuropsychologia.2003.12.016

Mayer, M. (1969). *Frog, Where Are You?* New York, NY: Penguin Group Publishers.

Miller, J., and Iglesias, A. (2008). *Systematic Analysis of Language Transcripts, Student Version 2008* [*Computer Software*]. Middleton, WI: SALT Software.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moriguchi, Y., and Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6017–6021. doi: 10.1073/pnas.0809747106

Mozeiko, J., Le, K., Coelho, C., Krueger, F., and Grafman, J. (2011). The relationship of story grammar and executive function following TBI. *Aphasiology* 25, 826–835. doi: 10.1080/02687038.2010.543983

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends Cogn. Sci.* 5, 309–315. doi: 10.1016/S1364-6613(00)01682-X

Munakata, Y. (2004). Computational cognitive neuroscience of early memory development. *Dev. Rev.* 24, 133–153. doi: 10.1016/j.dr.2003.09.008

NICHD Early Child Care Research Network. (2003). Do children's attention processes mediate the link between family predictors and school readiness?. *Dev. Psychol.* 39, 581–593. doi: 10.1037/0012-1649.39.3.581

Nicolopoulou, A., and Richner, E. (2007). From actors to agents to persons: the development of character representation in young children's narratives. *Child Dev.* 78, 412–429. doi: 10.1111/j.1467-8624.2007.01006.x

Perner, J., and Lang, B. (1999). Development of theory of mind and executive control. *Trends Cogn. Sci.* 3, 337–344. doi: 10.1016/S1364-6613(99)01362-5

Perner, J., and Lang, B. (2002). What causes 3−year−olds' difficulty on the dimensional change card sorting task? *Infant Child Dev.* 11, 93–105. doi: 10.1002/icd.299

Perner, J., Lang, B., and Kloo, D. (2002). Theory of mind and self−control: more than a common problem of inhibition. *Child Dev.* 73, 752–767. doi: 10.1111/1467-8624.00436

Posner, M. I., Rothbart, M. K., Sheese, B. E., and Tang, Y. (2007). The anterior cingulate gyrus and the mechanism of self-regulation. *Cogn. Affect. Behav. Neurosci.* 7, 391–395. doi: 10.3758/CABN.7.4.391

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Schneider, W., Lockl, K., and Fernandez, O. (2006). "Interrelationships among theory of mind, executive control, language development, and working memory in young children: a longitudinal analysis," in *Young Children's Cognitive Development,* eds W. Schneider, R. Schumann-Hengsteler, and B. Sodian (London: Psychology Press), 259–284.

Shapiro, L., and Hudson, J. (1991). Tell me a make-believe story: coherence and cohesion in young children's picture-elicited narratives. *Dev. Psychol.* 27, 960–974. doi: 10.1037/0012-1649.27.6.960

Smith, A. B., Taylor, E., Brammer, M., and Rubia, K. (2004). Neural correlates of switching set as measured in fast, event−related functional magnetic resonance imaging. *Hum. Brain Mapp.* 21, 247–256. doi: 10.1002/hbm.20007

Soveri, A., Rodriquez-Fornells, A., and Laine, M. (2011). "Is there a relation between language switching and executive functions in bilingualism? Introducing a within-groups analysis approach," in *Bilingualism and Cognitive Control,* ed J. F. Kroll (Lausanne: Frontiers in Psychology), 138–143.

Stelzer, F., Mazzoni, C. C., and Cervigni, M. A. (2014). Cognitive models of executive functions development. Methodological limitations and theoretical challenges. *Anales de psicología* 30, 329–336. doi: 10.6018/analesps.30.1.139251

Sulzby, E. (1985). Children's emergent reading of favorite storybooks: a developmental study. *Read. Res. Q.* 20, 458–481. doi: 10.1598/RRQ.20.4.4

Troiani, V., Fernández-Seara, M. A., Wang, Z., Detre, J. A., Ash, S., and Grossman, M. (2008). Narrative speech production: an fMRI study using continuous arterial spin labeling. *Neuroimage* 40, 932–939. doi: 10.1016/j.neuroimage.2007.12.002

Valencia, S. W., and Sulzby, E. (1991). Assessment of emergent literacy: storybook reading. *Read. Teach.* 44, 498–500.

Van der Ven, S. H., Kroesbergen, E. H., Boom, J., and Leseman, P. P. (2013). The structure of executive functions in children: a closer examination of inhibition, shifting, and updating. *Br. J. Dev. Psychol.* 31, 70–87. doi: 10.1111/j.2044-835X.2012.02079.x

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., and Weintraub, S. (2013). NIH toolbox cognitive function battery (CFB): Measuring executive function and attention. *Monogr. Soc. Res. Child Dev.* 78, 16–33. doi: 10.1111/mono.12032

Zelazo, P. D., and Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Curr. Dir. Psychol. Sci.* 7, 121–126.

Zelazo, P. D., and Jacques, S. (1996). Children's rule use: representation, reflection, and cognitive control. *Ann. Child Dev.* 12, 119–176.

Zelazo, P. D., Muller, U., Frye, D., and Marcovitch, S. (2003). The development of executive function in early childhood. *Monogr. Soc. Res. Child Dev.* 68, 1–124. doi: 10.1111/j.0037-976X.2003.00261.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 February 2014; accepted: 01 May 2014; published online: 20 May 2014.
*Citation: Friend M and Bates RP (2014) The union of narrative and executive function: different but complementary. Front. Psychol. **5**:469. doi: 10.3389/fpsyg.2014.00469*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Friend and Bates. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 11 April 2014
doi: 10.3389/fpsyg.2014.00291

Sharing and giving across adolescence: an experimental study examining the development of prosocial behavior

Berna Güroğlu[1]*, Wouter van den Bos[2] and Eveline A. Crone[1]

[1] *Institute of Psychology, Leiden University, Leiden, Netherlands*
[2] *Center for Adaptive Rationality, Max-Planck-Institute for Human Development, Berlin, Germany*

Edited by:
Philip D. Zelazo, University of Minnesota, USA

Reviewed by:
Valerie Kuhlmeier, Queen's University, Canada
Felix Warneken, Harvard University, USA
Tamar Kushnir, Cornell University, USA (in collaboration with Nadia Chernyak)

****Correspondence:***
Berna Güroğlu, Institute of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, Netherlands
e-mail: bguroglu@fsw.leidenuniv.nl

In this study we use economic exchange games to examine the development of prosocial behavior in the form of sharing and giving in social interactions with peers across adolescence. Participants from four age groups (9-, 12-, 15-, and 18-year-olds, total $N = 119$) played three types of distribution games and the Trust game with four different interaction partners: friends, antagonists, neutral classmates, and anonymous peers. Nine- and 12-year-olds showed similar levels of prosocial behavior to all interaction partners, whereas older adolescents showed increasing differentiation in prosocial behavior depending on the relation with peers, with most prosocial behavior toward friends. The age related increase in non-costly prosocial behavior toward friends was mediated by self-reported perspective-taking skills. Current findings extend existing evidence on the developmental patterns of fairness considerations from childhood into late adolescence. Together, we show that adolescents are increasingly better at incorporating social context into decision-making. Our findings further highlight the role of friendships as a significant social context for the development of prosocial behavior in early adolescence.

Keywords: friendship, prosocial behavior, fairness, trust, reciprocity, adolescence, peer relationships

INTRODUCTION

Prosocial behavior, defined as voluntary behavior intended to benefit others (Eisenberg et al., 2006), plays a key role in social interactions. Displays of prosocial behavior strengthen future ties between individuals and are crucial for the formation and continuation of relationships (Fehr et al., 2002). Although most studies have examined interactions with anonymous others, the majority of our social interactions are with people we know. Social behavior depends heavily on the relation we have with our interaction partners, such that prosocial behavior (including displays of fairness, trust, and reciprocity) is employed based on past experiences with the interaction partner and the prospect of future interactions (Burnham et al., 2000; Delgado et al., 2005; van den Bos et al., 2011a). This raises the question how prosocial behavior in these anonymous games reflects, or differs from, social behavior toward familiar peers. From a developmental perspective, the role of peer relationships in social interactions is an intriguing question given that with age there is a growing focus on peers, and that by adolescence individuals spend the majority of their time with them (Brown, 2004). As such, the peer group has been identified as one of the most significant developmental contexts with profound effects on the development of prosocial behavior (Carlo et al., 1999). This paper aimed to specifically examine the development of sharing and giving as observed in fairness- and trust-related social decisions when interacting with peers.

Prosocial behavior in the form of sharing and giving typically involves making decisions involving consequences for others and is based on comparisons of outcomes for self and others. These behaviors have been examined using different sorts of allocation games, which typically involve the distribution of resources between two players (Rilling and Sanfey, 2011). In these games with varying rules, the first player (i.e., the proposer) is typically asked to make a decision (i.e., an offer) on how to divide the stake between him/herself and a second player (i.e., the responder). In the current study, we focused on two types of allocation games that are specifically well-suited to study prosocial behavior in the form of sharing and giving.

The first type involves a set of allocation games developed to study fairness considerations, which refer to the direct comparison of outcomes for self and other (Fehr et al., 2008). In these games, the players are asked to choose between a fair distribution of goods (e.g., coins) with equal pay-offs to both players and an alternative unfair distribution that might be advantageous or disadvantageous for the self. Using these games with differing alternative distributions it is possible to systematically examine the role of costs to the self in sharing and giving. Prosocial responding assessed by such experimental paradigms is already shown in two and a half-year-old children, whose behavior is not contingent on prosocial or selfish behavior of their interaction partners (Sebastián-Enesco et al., 2013). Already by 3 years of age, children have an understanding of the fairness norm and that others expect them to share equally (Smith et al., 2013). Fehr et al. (2008) demonstrated that there is an increase in the preference for fair (or equal) splits between age 3 and 8 years. This finding is in line with prior studies with varying allocation paradigms showing that equity preferences increase across early childhood, even at the cost of throwing away resources (Blake and McAuliffe, 2011; Shaw and Olson, 2012). Using a similar choice-card task where participants could choose between different allocations of points for themselves and friends, Berndt (1985) has also shown

an age related increase in preferences for equal distributions over competition between 10 and 14 years of age. Recently, Steinbeis and Singer (2013) have provided further support for the developmental pattern of age related increase in these equity (fairness) preferences between the age of 7 and 13 years.

Despite the general trend of age related increase in fairness preferences (as assessed by relative number of fair/equal splits chosen) across these different games, differences in these preferences based on context have also been demonstrated. For example, both Fehr et al. (2008) and Steinbeis and Singer (2013) have shown that the preference for equal distributions was lower when they were costly than when they did not incur costs for the self. Further, age differences in choosing fair distributions were less pronounced when these choices were not costly than when they incurred costs. These findings suggest that the preference for fairness is dependent on the context regarding available alternatives. In the current study, we aimed to further examine these context effects in fairness related prosocial behavior in relation to different interaction partners.

A second sort of allocation paradigm suitable for examining sharing and giving is the Trust game (Berg et al., 1995). Trust behavior refers to decisions that favor other-regarding outcomes with the hope of future cooperation and self-gain (Larson, 1992). Reciprocity, such as returning a favor, refers to mutual exchange and is crucial for maintaining positive interactions (Lahno, 1995). In the Trust game, a first player can trust a second player to divide a stake, and the second player's reciprocity is an index for returning the favor initiated by the first player. In this sense, the trust choice assesses the extent of willingness to share and reciprocity assesses giving back. Interestingly, in these studies prosocial behavior, as indexed by level of trust and reciprocity, is even observed in one-shot social interactions with anonymous others where there is no prospect of future interactions between the two players. Developmental studies with the Trust game suggest that there are age related increases in trust and reciprocity toward anonymous others (Sutter and Kocher, 2007; van den Bos et al., 2010).

A social information processing approach has proven valuable to understanding the development of prosocial behavior. Prosocial young adolescents are shown to hold benign attributions, prefer to maintain a positive relationship with aggressive provocateurs, and show less negative emotionality in interactions (Nelson and Crick, 1999). Several studies have specifically focused on the role of dyadic characteristics in social behavior, showing that social-cognitive evaluations and behaviors are specific for interaction partners (Card and Hodges, 2007). Accordingly, interaction partners can evoke emotions that influence perception as well as processing of information, which together determine the behavioral output in context. For example, 4-year-olds attribute different emotions to the target depending on whether the target is a friend or a neutral classmate and are also more ready to help the target if the target is a friend. In adolescence, hostile attribution errors toward a specific peer are related to reactive aggression perceived from that peer (Hubbard et al., 2001; see also Ray and Cohen, 1997; Peets et al., 2007; Nummenmaa et al., 2008). In the current study we took a dyadic perspective in examining social behavior in the peer relationship context and we specifically expected that peer relations crucially influence displays of prosocial behavior.

In the current study we investigated how prosocial behavior is influenced by peer relationships by combining allocation games with sociometric mapping of relationships within across a wide age range of 9 to 18 years. Participants played a set of three allocation games (Fehr et al., 2008) and a Trust game (Berg et al., 1995) with four interaction partners: friends, antagonists, neutral peers, and anonymous peers. Based on prior studies using one-shot interactions (Sutter, 2007; Güroğlu et al., 2009b; van den Bos et al., 2010), we expected that in the current study participants would show increasing levels of prosocial behavior (defined as choices maximizing other's outcome) with increasing age.

A number of studies with varying methodology, paradigms, and measures have shown that children treat friends and non-friends differently. There is evidence for this differential treatment of in-group members (classmates/friends) vs. out-group members (anonymous peers/strangers) already by age three or four (Costin and Jones, 1992; Fehr et al., 2008; Moore, 2009), also when children are interacting with a doll protagonist (Olson and Spelke, 2008). Similarly, 3-year-olds are shown to share equally with collaborators (Warneken et al., 2011) and 5-year-old children display strong ingroup preferences with random group assignment and lack of a competitive context, both in terms of implicit and explicit attitudes, as well as resource allocation (Dunham et al., 2011). Some studies show a further differentiation between familiar peers. Examining reward allocations and helping behavior, Berndt (1985) has shown that young adolescents treat interaction partners differentially: adolescents were more generous and helping toward friends than toward neutral classmates. Similarly, Buhrmester et al. (1992) have shown that children and adolescents share more with friends than with neutral peers and share least with disliked peers; Amato (1990) has also shown that young adults help friends more than they help strangers. In the current study, we aimed to move beyond a dichotomous exploration of ingroup vs. outgroup members and examine peer relationships with varying valence (positive, negative, and neutral) and compared to unfamiliar peers. Furthermore, the majority of these previous studies have examined early childhood, whereas less is known about the changes in social decision-making across adolescence. In the current study, we focus on a broad age range across middle childhood and adolescence (9- to 18-year olds) where we can assess peer relationships in a structured environment, i.e., the classroom, using the same methodology, i.e., sociometric nominations. We expected that prosocial behavior would be moderated by the interaction partner, where participants were expected to display highest levels of prosocial behavior toward friends and lowest levels toward antagonists. We also expected this differentiation to be modulated by the specific allocation game.

It has further been shown that young adolescents become more relationship-focused with age, as indicated by more relational attributions to provocations from peers (Nelson and Crick, 1999). This is in line with the theoretical perspectives in changes in interpersonal interactions in general, and in friendships in particular, across adolescence (Selman, 1980). The development of cognitive skills and perspective-taking across adolescence are

central to Selman's theory of interpersonal growth. Previous findings showing that older adolescents are increasingly better able to incorporate context related information into their decision-making process are further in line with these theoretical perspectives (Güroğlu et al., 2009a,b). Along similar lines, Berndt (1985) has shown that 14-year-olds differentiate more between friends and neutral classmates than 10- and 12-year-olds in displays of generosity. Such findings are also supported by studies examining the development of friendships. Around late childhood and early adolescence there is a specific increase in prosocial behavior such as helping and sharing as well as a concern for equality in interactions with friends (Youniss, 1980; Berndt, 1981; Furman and Bierman, 1984). This age related difference on the increasing specificity of friends was expected to reflect in age related differences in prosocial behavior toward friends in the current study. Taken together, we expected the moderation by interaction partner in prosocial behavior levels to be more pronounced for older participants than for younger ones.

One of the mechanisms that may account for developmental differences in prosocial behavior is the ability to take the other player's perspective. From a developmental perspective, the cognitive ability of role taking has implications for the development of altruistic motivation and behavior (Hoffman, 1975). Experimental studies in children as young as 3–4 years old show links between theory of mind skills and future-oriented prosocial behavior (Moore et al., 1998). A positive relation between prosocial behavior and perspective-taking skills has long been established (Eisenberg and Miller, 1987; Eisenberg et al., 1991; Carlo and Randall, 2002). It has been suggested that the components that are related to the consistency of prosocial behavior across time are related to, besides temperamental/genetic predispositions, inhibitory control and "other-orientation" (Eisenberg et al., 1999). This component of "other-orientation" is tapped by the cognitive ability to take others' perspectives and incorporate these perspectives into decision-making, which continues to develop into late adolescence (Dumontheil et al., 2009). The development of this ability of perspective-taking in social settings has been suggested to be a mediator of the development of prosocial behavior with increasing age (Iannotti, 1985). In prior studies we demonstrated the role of perspective taking by correlating the self-report index of the Interpersonal Reactivity Index (IRI, Davis, 1983) with prosocial behavior (Overgaauw et al., 2012), as well as a relation between affective perspective taking and prosocial behavior in the form of costly compensation of victims (Will et al., 2013). In the current study, we tested for the mediating role of perspective-taking skills in the development of prosocial behavior. We expected that the age related increase in prosocial behavior in both the set of allocation games and the Trust game would be more pronounced for individuals with higher levels of self-reported perspective taking.

METHODS

PARTICIPANTS

A total of 125 participants took part in the study. The majority of the participants (90.4%) were Dutch, 2.4% was of Moroccan decent and 4.0% had another ethnic background; ethnic background information of four participants (3.2%) was missing. In order to control for the role of a general cognitive capacity, we assessed and controlled for IQ in our analyses. The pen-and-paper version of the Raven Standard Progressive Matrices (SPM) (Carpenter et al., 1990) was administered to assess an estimate of the participant's intelligence quotient (IQ). Due to time restrictions Raven scores of four participants were missing. After removing six outliers with IQ two standard deviations higher than the mean, estimate scores on IQ ranged between 94 and 130; the mean was 114.17 ($SD = 9.37$). The remaining 119 participants consisted of: 9-year-olds (M age = 9.27 years, $SD = 0.53$, 15 boys and 16 girls), 12-year-olds (M age = 11.89 years, $SD = 0.64$, 18 boys and 14 girls), 15-year-olds (M age = 15.07 years, $SD = 050$, 13 boys and 12 girls), and 18-year-olds (M age = 17.95 years, $SD = 0.54$, 8 boys and 23 girls). There were no differences in the gender distribution across age groups [$\chi^2_{(3)} = 6.87$, $p = 0.08$]. Thus, the sample sizes per age group ranged between 25 and 31, which is comparable to previous studies employing similar experimental designs (Fehr et al., 2008; Steinbeis and Singer, 2013).

There was a significant difference in IQ scores between the age groups [$F_{(3,\ 111)} = 5.62$, $p = 0.001$]. Tukey *post-hoc* tests showed that 18-year-olds had higher IQ ($M = 119.65$, $SD = 8.37$) than all other younger age groups ($M = 112.50$, $SD = 10.37$, $M = 112.66$, $SD = 8.54$, and $M = 110.91$, $SD = 7.61$, respectively for 9-, 12-, and 15-year-olds). Therefore, all analyses were run including IQ as a covariate; as suggested by Delaney and Maxwell (1981) the covariate was mean centered for ANCOVA analyses in a repeated measures design. There were no main effects of or interactions with IQ scores in any of the analyses reported below.

MATERIALS

Peer relationships

Friendship and antipathy relationships were identified based on sociometric nominations, and neutral peer relationships were based on peer ratings. Participants were provided with a numbered list of all classmates and were asked nominate up to five classmates for the questions "Who are your friends?" and "Who do you not like at all?" Mutual nominations on these items were used to identify friendship and antipathy relationships (i.e., positive and negative peer relationships), respectively (Güroğlu et al., 2007, 2009a). In addition, participants were asked to rate how much they liked each classmate on a scale ranging from (1) "do not like at all" to (3) "neither like nor dislike" to (5) "like very much." Classmates who mutually gave a neutral rating (3) for one another were identified as neutral peer relationships.

Perspective-taking

Perspective-taking was measured by the Perspective-taking subscale of the IRI (Davis, 1983). This measure of perspective-taking was included because it (i) assesses the tendency to spontaneously adopt the psychological point of view of others (rather than e.g., a spatial point of view), (ii) assesses cognitive empathy skills (rather than e.g., affective empathy), (iii) is related to measures of interpersonal functioning, and (iv) is suitable for the broad age range of 9 to 18 years old. The perspective-taking subscale consisted of 6 items (e.g., "I try to look at everybody's side of a disagreement before I make a decision") answered on a 5-point Likert scale

ranging from (1) not true at all to (5) completely true. We used an adolescent version of the IRI, where items have been adapted for the youngest age group in the study. The scale had moderate reliability (Cronbach's alpha 0.68).

Fairness-related prosocial behavior

A set of three allocation games were used to assess prosocial behavior related to fairness considerations (Fehr et al., 2008). Participants played these games on the computer where they were asked to distribute coins between themselves and their interaction partner by choosing one of the two preset distributions. One of the two options in each game was a fair distribution of coins with one coin for the self and one coin for the interaction partner [i.e., (1/1) distribution]. The alternative option varied between the three games, yielding three games: (i) the *Costly prosocial* game where the alternative option was two coins for self and zero coins for the other [i.e., (2/0) distribution], (ii) the *Non-costly prosocial* game where the alternative option was one coin for self and zero coins for the other [i.e., (1/0) distribution], (iii) the *Disadvantageous prosocial* game where the alternative option was one coin for the self and two coins for the other [i.e., (1/2) distribution] (see **Figure 1**). The dependent variable was the frequency of *prosocial* (i.e., not self-focused) choices [i.e., (1/1) distribution in the *Non-costly* and *Costly prosocial* games and (1/2) distribution in the *Disadvantageous prosocial* game] and was calculated separately per game and interaction partner.

Participants played a total of 48 trials of games in randomized order. The location of the fair distribution (1/1) was counterbalanced across trials. All three games were played four times with each of the four interaction partners (friends, antipathies, neutral, and anonymous peers). In order to render the games less repetitive and keep the participants engaged in these multiple trials, we used the following design: Participants were told that each round of the game would be played with one of the four groups that were predetermined by the researchers. They were explained that the peers in three of the four groups would be randomly chosen classmates and the fourth group would be anonymous same gender and age peers from another school. In fact, peers from the three groups with classmates were not randomly chosen classmates. Each of the three groups contained either friends, neutral peers, or antagonists identified based on the sociometric nominations and ratings obtained during the first data collection. In each group, there were one, two or three players.

Care was taken to present all four groups in a neutral manner so that participants would not be biased toward one group or another. To accomplish this, each group was randomly given one of the following neutral names: group Bike, group Car, group Airplane, and group Train. Participants were given lists of players in each group and were given ca. 5 min to study the group members. During each trial of the game, the list of players within a group was presented on the left side of the screen (see **Figure 1**). Each group was randomly assigned to the group of friends, antagonists, neutral classmates, and anonymous interaction partners.

Participants were told that they would play each trial with a single individual interaction partner from the group they were playing with but they would not know *exactly* with whom. This was done so that there would be no strategies for multiple distributions. It was further explained that the computer would keep track of their interaction partners in each trial in order calculate everyone's earnings, which would be paid out at the end of all trials. Each trial started with a fixation cross (1 s), followed by a screen presenting the group they are playing with (left panel) and the set of alternatives they could choose from. Participants had 5 s to respond by pressing a keyboard key. If they failed to respond within 5 s, a screen with "Too late!" was presented for 1 s. Upon response, their choice was encircled in red for 2 s and subsequently they were presented with the following trial. Completion of this task took about 2 min on average. Participants played six practice trials with the computer before the actual games started.

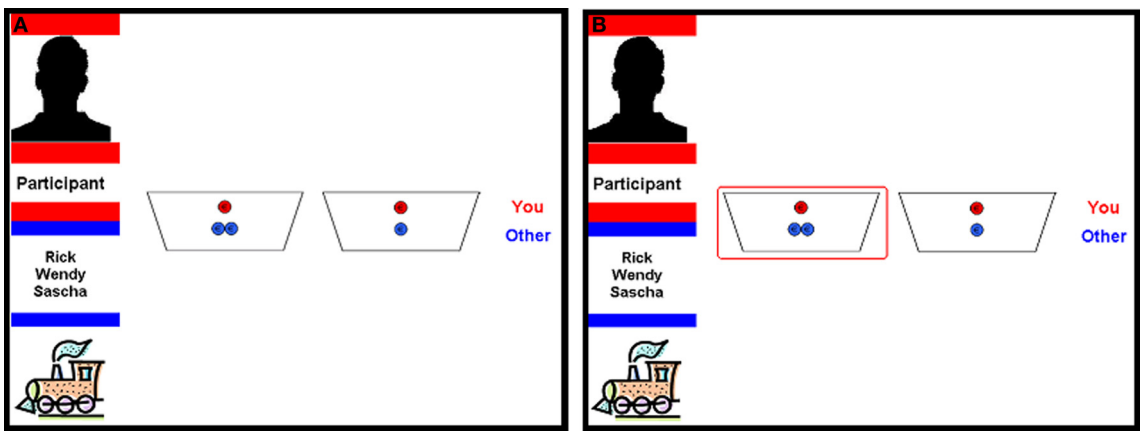


FIGURE 1 | Visual display for the allocation games. (A) Two offers, each containing red and blue coins, indicate the share for the proposer and the interaction partner, respectively (here depicted *Disadvantageous prosocial* game 1/2 vs. 1/1). The left top panel displays the name of the proposer in red (here "Participant"). The left bottom panel displays the group (here group "train") in the current trial and the names of the players in this group (here "Rick, Wendy, and Sascha"). **(B)** The red encircled option indicates the offer made by the participant.

Trust-related prosocial behavior

A single round of an adaptation of the Trust game (Berg et al., 1995) was used to assess trust and reciprocity in social interactions. Participants played the Trust game on paper once as the first player (investor) and once as the second player (trustee) with each of the four types of interaction partners (i.e., 8 rounds in total). The four interaction partners were presented in four groups in the same way as for the allocation games explained above. The starting stake was 10 coins and the first player could choose between two options: an equal distribution of 5 coins for self and 5 coins for the trustee, or letting the trustee decide (i.e., trust). In the latter case, the stake was doubled and the trustee had two options: give 10 coins each (i.e., reciprocate) or give nothing to the investor and take 20 coins for him/herself (i.e., defect). The options for the second player were visible to the first player from the start (see **Figures 2A** and **2B** respectively for participant as investor and trustee). The dependent variable was the prosocial (i.e., not self-focused) choices made by the players and was coded in the following way: as the investor, the trust option was coded as 1 and no-trust option as 0; as the trustee, the reciprocate option was coded as 1 and defect option as 0. Average frequency of trust and reciprocity were calculated per age group and interaction partner.

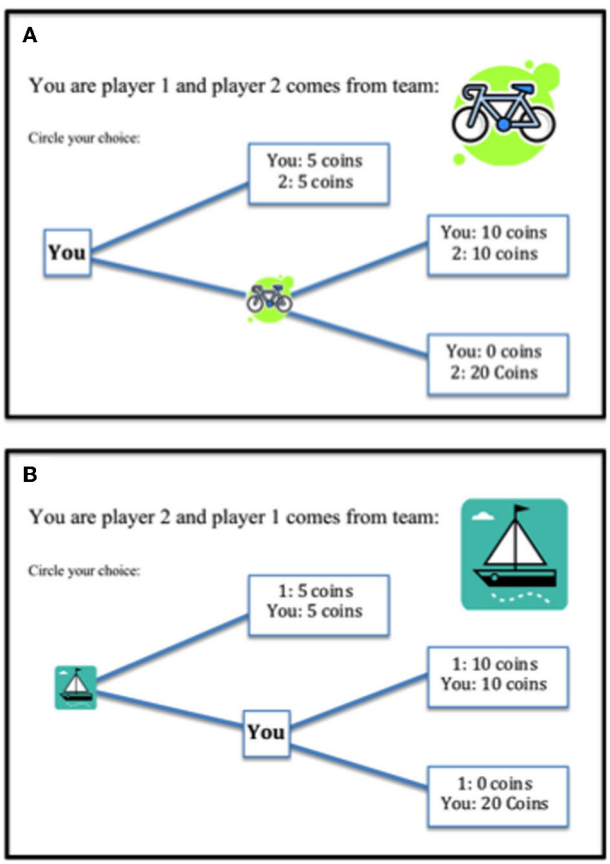


FIGURE 2 | Visual display for the Trust game. (A) Participant is the investor (here "you"), the interaction partner is the trustee (here group "boot"). **(B)** Participant is the trustee (here "you"), the interaction partner is the investor (here group "bicycle").

PROCEDURE

Two elementary schools and one high school agreed to take part in the study. After consent was obtained from school authorities, informed consent was obtained from parents and teachers. The first part of the data collection was carried out in classrooms where participants filled out several questionnaires, including sociometric nominations and ratings for all classmates, perspective-taking scale and the pen-and-paper version of the Raven's SPM. This session lasted about 45 min. Approximately 1 week later at a second data collection day, computer tasks were presented on individual laptops with 15-inch monitors in a separate room. In groups of four at a time, participants completed the allocation tasks on the computer and the Trust game on paper. Care was taken that all instructions were clear. Previous studies have successfully employed similar experimental designs with computer based allocation and gambling games in the age groups assessed here (van Leijenhorst et al., 2008; Güroğlu et al., 2009a,b; van den Bos et al., 2010). This session lasted for about 30 min. At both data collection points participants were explained that their participation was voluntary and were ensured that their responses would remain anonymous. In order to further assure anonymity, we also emphasized during the second data collection point that the computer tasks were not online interactions and that classmates could not see the participants' responses. We also took care to place individual laptop computers facing away from each other so that it was not possible for the participants to view each other others' responses.

Participants were told that the coins in the allocation tasks were valuable. It was explained that after all participants completed the allocation tasks and data collection was completed each participant would be paid a randomly chosen number of trials. It was emphasized that their decisions would determine the earnings for themselves as well as for their interaction partners. After data collection was completed, in agreement with the schools and parents all participants were paid a fixed amount of 3 euros (∼5 US dollars) each. This procedure was approved by the local ethics committee.

RESULTS

MANIPULATION CHECK

At the end of the second session participants were asked to make a list of players in each group and were asked to indicate what they thought of each group (except for the group with anonymous players). This was assessed as a manipulation check to ensure that the participants paid attention to the group members and that they distinguished between the three groups of classmates each containing friends, antagonists, and neutral classmates in terms of likeability. Percentage of correct recall for the players in each group was high ($M = 82\%$, $SD = 20\%$). Fifteen-year-olds recalled significantly more players than 12-year-olds [$M = 91$ and 75%, respectively; $F_{(3,\ 104)} = 3.53$, $p = 0.02$]. Participants recalled players from the friend group (91%) more often than players from the antagonist (82%) and neutral peer groups (79%)

[$F_{(1,\ 91)} = 13.0, p < 0.001$ and $F_{(1,\ 95)} = 16.8, p < 0.001$, respectively]. Open-ended questions on what the participants thought of each group were recorded on a five-point scale ranging from (1) very negative to (5) very positive. Participants rated the friend group ($M = 4.98$, $SD = 0.06$) more positive than the neutral group [$M = 3.84$, $SD = 0.16$; $F_{(1,\ 99)} = 47.7$, $p < 0.001$], which was rated more positively than the antagonist group [$M = 3.26$, $SD = 0.23$; $F_{(1,\ 94)} = 4.90$, $p = 0.03$]. This manipulation check confirmed our expectation that the participants differentiated between different groups in terms of their relationships with them. The ratings for each group did not differ across the age groups [$F_{(3,\ 88)} = 0.92$, $p = 0.44$, $\eta_P^2 = 0.03$]; there was also no age group × group interaction in the ratings [$F_{(5.27,\ 154.64)} = 1.19$, $p = 0.32$, $\eta_P^2 = 0.04$].

DESCRIPTIVES

Peer relationships

The mean number of mutual friendships, antipathies, and neutral relationships were 2.75 ($SD = 1.62$), 0.36 ($SD = 0.76$), and 5.26 ($SD = 4.51$), respectively. Univariate analyses of variance (ANOVA) with age group as the between subjects factors yielded a main effect of age for number of friendships and antipathies [$F_{(3,\ 113)} = 2.75$, $p = 0.05$, $\eta_P^2 = 0.07$ and $F_{(3,\ 113)} = 3.54$, $p = 0.02$, $\eta_P^2 = 0.09$, respectively]. There were more friendships in 15-year-olds ($M = 3.43$, $SD = 1.50$) than in 18-year-olds ($M = 2.26$, $SD = 1.32$). Nine-year-olds ($M = 0.71$, $SD = 1.07$) had more antipathy relationships than 18-year-olds ($M = 0.13$, $SD = 0.34$).

Perspective-taking

The perspective-taking scores ranged from 1.17 to 4.83 with a mean of 3.32 ($SD = 0.63$). There was a significant correlation between perspective-taking and age [$r_{(117)} = 0.35$, $p < 0.001$] and between perspective-taking and IQ [$r_{(113)} = 0.23$, $p = 0.02$]; the correlation between age and perspective-taking remained significant when controlling for IQ [partial $r_{(110)} = 0.34$, $p < 0.001$].

PROSOCIAL BEHAVIOR IN FAIRNESS CONSIDERATIONS

A repeated measures analysis of variance with Age group (four levels: 9-, 12-, 15-, and 18-year-ols) as the between subject factors and Relationship type (four levels: friends, antipathies, neutral peers, and anonymous peers) as the within subject factor was conducted for frequency of prosocial offers made in each of the three games[1]. All analyses where the Mauchly's test indicated a violation of the assumption of sphericity, the Huyn-Feldt correction is reported.

In the *Non-costly prosocial* game (see **Figure 3A**), participants chose the prosocial offer [i.e., (1/1) distribution] on 52% of the trials ($SD = 29\%$). The main effect of Age group was not significant [$F_{(3,\ 103)} = 2.34$, $p = 0.08$, $\eta_p^2 = 0.06$]. There was a main effect of Relationship type [$F_{(3,\ 309)} = 16.5$, $p < 0.001$, $\eta_p^2 = 0.14$]: prosocial behavior was higher for friends than for neutral peers [$F_{(1,\ 106)} = 4.58$, $p = 0.04$, $\eta_p^2 = 0.04$], which was again higher than for antagonists [$F_{(1,\ 106)} = 13.6$, $p < 0.001$, $\eta_p^2 = 0.11$]. Prosocial behavior toward antagonists and anonymous peers did not differ [$F_{(1,\ 106)} = 0.04$, $p = 0.85$]. This main effect was qualified by an Age group × Relationship type interaction [$F_{(9,\ 309)} = 2.63$, $p = 0.006$, $\eta_p^2 = 0.07$]. Nine- and 12-year-olds did not differ in their frequency of (1/1) offers across the four interaction partners [overall $M = 57$ and 45%, $SD = 26\%$ and 28%, $F_{(2.41,\ 48.3)} = 0.69$, $p = 0.56$, $\eta_p^2 = 0.02$ and $F_{(3,\ 81)} = 1.76$, $p = 0.16$, $\eta_p^2 = 0.06$, respectively]. In contrast, 15- and 18-year-olds differentiated in their responses toward the other players [$F_{(2.41,\ 48.3)} = 5.26$, $p = 0.006$, $\eta_p^2 = 0.21$ and $F_{(3,\ 75)} = 6.22$, $p = 0.001$, $\eta_p^2 = 0.20$, respectively]. Tukey *post-hoc* tests indicated that 15- and 18-year-olds were more prosocial toward friends ($M = 63$ and 82%, respectively) than toward antagonists and anonymous peers ($M = 37$ and 46% anonymous peers, and $M = 34$ and 46% antagonists, respectively for 15- and 18-year-olds; all $F > 6.21, p < 0.02$). Further, both 15- and 18-year-olds displayed more prosocial behavior toward the neutral peers ($M = 53$ and 69%, respectively) than toward antagonists ($M = 34$ and 46%, respectively; all $F > 4.90, p < 0.04$).

In the *Costly prosocial* game (see **Figure 3B**), participants chose the fair (1/1) distribution on approximately 50% of the trials ($SD = 29\%$). The main effect of Age group was not significant [$F_{(3,\ 103)} = 0.9$, $p = 0.47$, $\eta_p^2 = 0.03$]. There was a main effect of Relationship type [$F_{(2.87,\ 296)} = 18.7$, $p < 0.001$, $\eta_p^2 = 0.15$]. As in the *Non-costly prosocial* game, prosocial behavior was again higher for friends than for neutral peers [$F_{(1,\ 106)} = 9.07$, $p = 0.003$, $\eta_p^2 = 0.08$], which was again higher than for antagonists [$F_{(1,\ 106)} = 13$, $p < 0.001$, $\eta_p^2 = 0.11$]; prosocial behavior toward antagonists and anonymous peers again did not differ [$F_{(1,\ 106)} = 0.28$, $p = 0.60$]. This interaction was qualified by an Age group × Relationship type interaction [$F_{(8.62,\ 296)} = 2.33$, $p = 0.02$, $\eta_p^2 = 0.06$]. Again, 9- and 12-year-olds did not differ in their frequency of prosocial offers across their interaction partners [overall $M = 55$ and 51%, $SD = 29$ and 27%; $F_{(2.58,\ 72.2)} = 0.92, p = 0.42$ and $F_{(3,\ 81)} = 2.53, p = 0.06$, respectively]. For the other two age groups, a differentiation was observed [$F_{(3,\ 60)} = 3.53$, $p = 0.02$, $\eta_p^2 = 0.15$ and $F_{(3,\ 75)} = 4.84$, $p = 0.004$, $\eta_p^2 = 0.16$, for 15- and 18-year-olds, respectively]: participants displayed more prosocial behavior toward friends (15-year olds $M = 57\%$; 18-year-olds $M = 75\%$) than toward anonymous peers (15-year-olds $M = 28\%$; 18-year-olds $M = 34\%$; all $F > 4.99$, $p < 0.04$) and antagonists (15-year-olds $M = 34\%$; 18-year-olds $M = 39\%$; all $F > 7.58$, $p < 0.01$). Furthermore, 18-year-olds were also more prosocial toward their friends than toward neutral peers [$M = 51\%$, $F_{(1,\ 25)} = 10.2$, $p = 0.004$, $\eta_p^2 = 0.29$].

Finally, in the *Disadvantageous prosocial* game (see **Figure 3C**), the prosocial (1/2) distribution was chosen on approximately one-third of the trials ($M = 32\%$, $SD = 26\%$). The main effect of Age group was not significant [$F_{(3,\ 98)} = 2.56$, $p = 0.06$, $\eta_p^2 = 0.07$]. There was again a main effect of Relationship type [$F_{(2.57,\ 252)} = 26.2$, $p < 0.001$, $\eta_p^2 = 0.21$]. Prosocial behavior was higher for friends than for neutral peers [$F_{(1,\ 101)} > 20.7$, $p < 0.001$, $\eta_p^2 = 0.17$]. Behavior toward

[1] See Supplementary material for analyses comparing behavior across the three games per age group.

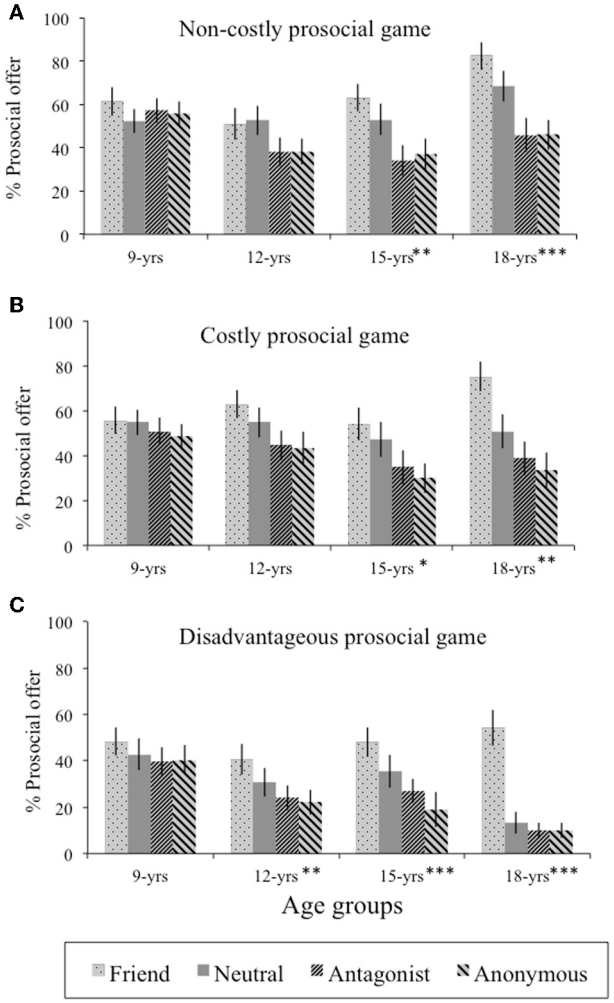


FIGURE 3 | Prosocial behavior in the allocation games. Mean frequency (%) and standard errors of prosocial offers [i.e., (1/1) distribution in the **(A)** *Non-costly prosocial game* and the **(B)** *Costly prosocial game* and (1/2) distribution in the **(C)** *Disadvantageous prosocial game*] are presented per interaction partner for the four age groups. Age differences are indicated by an asterisk (*). $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

neutral peers, antagonists, and anonymous peers did not differ significantly [$F_{(1,\ 101)} = 3.70$, $p = 0.06$, $\eta_p^2 = 0.04$]. There was also a significant Age group × Relationship type interaction in the *Disadvantageous prosocial* game [$F_{(7.71,\ 252)} = 2.36$, $p = 0.02$, $\eta_P^2 = 0.07$]. As in the *Non-costly prosocial* game and the *Costly prosocial* game, 9-year-olds did not differ in frequency of prosocial choices across interaction partners [overall $M = 43\%$, $SD = 26\%$, $F_{(1,\ 28)} = 1.50$, $p = 0.23$]. In contrast, 12-, 15-, and 18-year-olds were more prosocial toward their friends ($M = 41\%$, $M = 48\%$, and $M = 54\%$, respectively; all $F > 7.81$, $p < 0.01$) than toward antagonists ($M = 24\%$, $M = 27\%$, and $M = 10\%$, respectively) and anonymous peers ($M = 22\%$, $M = 19\%$, and $M = 10\%$, respectively; all $F > 9.90$, $p < 0.004$). Both 15- and 18-year-olds displayed also more prosocial behavior toward their friends than toward neutral peers ($M = 35\%$ and $M = 13\%$, respectively, all $F > 6.38$, $p < 0.02$).

PROSOCIAL BEHAVIOR IN TRUST AND RECIPROCITY CONSIDERATIONS

Two repeated measures analyses were conducted; one for trust and one for reciprocity choices with Age group as the between subjects factor and Relationship type as the within subject factor. For trust behavior (see **Figure 4A**), there was only a significant main effect of relationship [$F_{(3,\ 255)} = 37.7$, $p < 0.001$, $\eta_P^2 = 0.31$]. Participants trusted friends ($M = 72\%$, $SD = 45\%$) more often than other peers (all $F > 61.7$, $p < 0.001$). Trust displayed for antagonists ($M = 21\%$, $SD = 41\%$), anonymous ($M = 20\%$, $SD = 40\%$) and neutral peers ($M = 29\%$, $SD = 46\%$) did not differ from each other [$F_{(2,\ 170)} = 2.40$, $p = 0.09$]. There was no main effect of Age group or an interaction with Age group.

For reciprocity (see **Figure 4B**), there was a only main effect of Relationship, with higher reciprocity for friends than for other interaction partners [$F_{(3,\ 255)} = 31.7$, $p < 0.001$, $\eta^2 = 0.27$]. Mean reciprocity ranged between 83% (9-year-olds, $SD = 38\%$) and 100% (18-year-olds, $SD = 0\%$). We examined the reciprocity scores for the other three interaction partners separately for the four age groups. These analyses showed that 9-, and 12-year-olds did not differ in reciprocity toward antagonists, neutral, and anonymous peers (all $F < 2.80$, $p > 0.08$). In contrast, 15- and 18-year-olds showed higher reciprocity toward neutral peers ($M = 63\%$, $SD = 50\%$ and $M = 68\%$, $SD = 48\%$, respectively) than toward anonymous peers [$M = 26\%$, $SD = 45\%$ and $M = 20\%$, $SD = 41\%$; $F_{(1,\ 17)} = 8.01$, $p = 0.01$, $\eta^2 = 0.32$ and $F_{(1,\ 23)} = 9.50$, $p = 0.005$, $\eta_P^2 = 0.29$, respectively for 15- and 18-year olds].

MEDIATING ROLE OF PERSPECTIVE-TAKING

Next, we investigated the mediating role of perspective-taking in the link between age and prosocial behavior. For this purpose, we followed the mediator analysis and SPSS syntax provided by Preacher and Hayes (2004). This method tests whether an indirect effect (i.e., the path from age to prosocial behavior with perspective-taking as mediator) is significantly different from zero. Accordingly, we examined the coefficients for (*a*) the link between the independent variable (i.e., age) and the mediator (i.e., perspective-taking), and (*b*) the link between the mediator (i.e., perspective-taking). We used a bootstrapping technique with 10,000 iterations and computed the 95% confidence interval around the product term *a***b*. The mediation effect is significant if zero falls out of this confidence interval. Considering that the direct effect of age on prosocial behavior is a prerequisite for testing mediation, we focused our analyses on those dependent variables where we observed a significant correlation with age: prosocial behavior with friends and neutral peers in the *Non-costly prosocial* game [$r_{(105)} = 0.21$, $p = 0.03$ and $r_{(105)} = 0.19$, $p = 0.05$, respectively], prosocial behavior with anonymous peers in the *Costly prosocial* game [$r_{(105)} = -0.21$, $p = 0.03$], prosocial behavior with antagonists, neutral peers, and anonymous peers in the *Disadvantageous prosocial* game [$r_{(100)} = -0.29$, $p = 0.003$, $r_{(100)} = -0.27$, $p = 0.006$, and $r_{(100)} = -0.31$, $p = 0.002$, respectively], and reciprocity with anonymous peers [$r_{(86)} = -0.23$, $p = 0.03$].

A significant mediation effect was found only for the *Non-costly prosocial* game with friends and not for the other dependent variables. The 95% confidence interval for the indirect effect

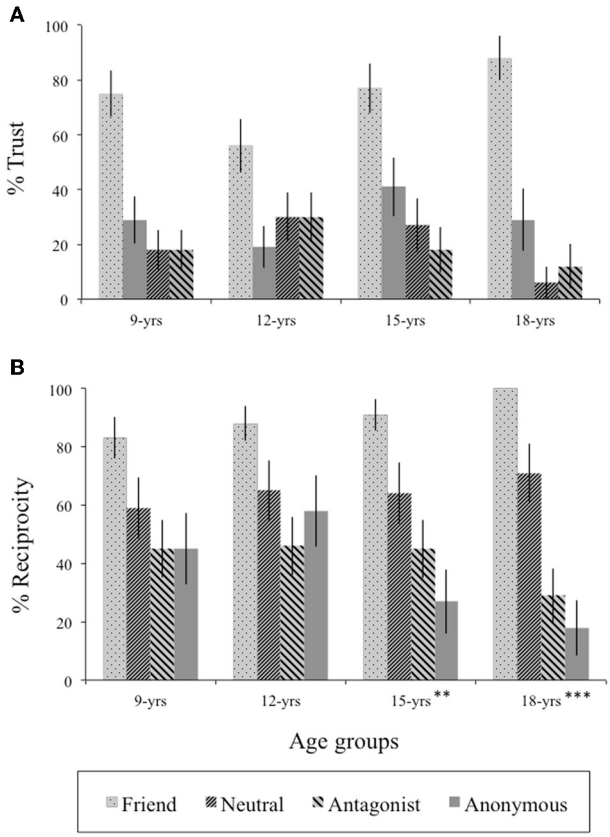


FIGURE 4 | Prosocial behavior (trust and reciprocity) in the Trust game. Mean frequency (%) and standard errors of **(A)** trust and **(B)** reciprocity choices in the Trust game are presented per interaction partner for the four age groups. Age differences are indicated by an asterisk (*). $^{**}p < 0.01$, $^{***}p < 0.001$.

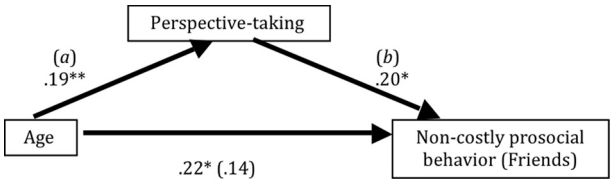


FIGURE 5 | Perspective-taking skills mediating the link between age and prosocial behavior. Figure depicts the results of the mediation analyses and shows that perspective-taking is a mediator between age and non-costly prosocial behavior toward friends. $^{**}p < 0.001$, $^{*}p < 0.05$.

ranged from 0.17 to 1.61, showing that perspective-taking mediates the direct link between age and prosocial behavior toward friends (see **Figure 5**). The direct effect of age on prosocial behavior was no longer significant when controlling for perspective-taking ($\beta = 0.14$), $t_{(111)} = 1.72$, $p = 0.09$.

DISCUSSION

The current study employed an experimental approach toward examining the development of prosocial behavior in social interactions with peers across adolescence. Our findings contribute to the existing literature examining context dependency of social behavior in three significant manners. First, we employed a variety of controlled experimental conditions examining forms of prosocial behavior such as costly and non-costly prosocial behavior, as well as trust and reciprocity, which provided us with different ways of assessing altruistic motivations aimed at maximizing outcomes for another person. Second, we examined behavior with four different interaction partners. Finally, we examined these processes across a wide age range from 9 to 18 years. More specifically, we demonstrated that 9- and 12-year-olds treated interaction partners similarly, whereas older adolescents' (15- and 18-years) prosocial behavior was significantly moderated by who their interaction partner was. Moreover, we demonstrated that perspective-taking skills mediated age related differences in prosocial behavior when interacting with friends.

DEVELOPMENT OF PROSOCIAL BEHAVIOR

We assessed prosocial behavior using a set of three allocation games: the *Non-costly prosocial* game, the *Costly prosocial* game, and the *Disadvantageous prosocial* game (Fehr et al., 2008). By presenting participants a dichotomous choice where one of the two options is always a (1/1) fair distribution, we were able to compare the preference for equal outcomes across different contexts. Three relevant processes need to be kept in mind in interpreting decision-making processes across these conditions: (1) a strong preference for equity, which would be indicated by equity choices (1/1) across games, (2) cost of choosing one distribution over the other in each game, and (3) payoff comparison for self vs. other, that is, whether the other gets more than self or not (Radke et al., 2012). A strong sense of equity requires participants to choose the (1/1) distribution regardless of context (i.e., game) with varying costs to the self.

Fehr et al. (2008) previously showed that prosocial behavior increases with age from 3 to 8 years, but that 8-year-olds have a stronger preference for equity, also when the alternative is a non-costly *and* prosocial distribution (i.e., in the *Disadvantageous prosocial* game). This result was replicated in the current study in adolescents. That is to say, overall levels of prosocial choices were lower in the *Disadvantageous prosocial* game than in the other two games, supporting context dependency of fairness considerations.

It is important to consider the current results in relation to previous findings. Although earlier findings have not been completely unanimous, several studies have shown no age differences in across 9 to 18 years in costly prosocial behavior assessed as fair allocations in a Dictator game (Gummerum et al., 2008; Güroğlu et al., 2009a,b). In the current study we also show that there are no age differences in the *Costly prosocial* game in interactions with classmates, whereas there is a slight age-related decline in fair allocations to anonymous others. Overall levels of prosocial behavior in the *Non-costly prosocial* game were somewhat lower than those reported by Fehr et al. (2008) for 7–8 year-olds (around 80%). However, Steinbeis and Singer (2013) reported equity choices in the *Non-costly prosocial* game to be around 15% for 7–8 year-olds, and around 60% for 11–13 year-olds, which is similar to our findings. In contrast, prosocial choices in the *Disadvantageous prosocial* game were higher in the current study than those reported previously, particularly for the youngest age group. As suggested by Steinbeis and Singer (2013), different

incentives used in these studies form a plausible explanation for these discrepancies. Furthermore, previous studies examined interactions with anonymous others in general, whereas the current study introduced different interaction partners. It is likely that such differences in the experimental design shape choices, where participants' decisions are influenced by the broad context in which different decisions are being made across interaction partners (for a similar discussion, see (Güroğlu et al., 2009a,b). Interestingly, percentages of prosocial choices in the *Costly prosocial* game were comparable across all three studies. Future studies could investigate whether non-costly prosocial behavior is more sensitive to context factors than costly prosocial behavior.

In addition, we showed that interaction partners significantly moderated the developmental patterns of prosocial behavior across ages 9 to 18. Specifically, there was an age related increase in non-costly prosocial behavior (i.e., in the *Non-costly prosocial* game), but only toward friends and neutral peers. Costly prosocial behavior decreased with age toward anonymous peers. Thus, participants are willing to incur costs for an equitable distribution, but with increasing age less so for unknown others. Finally, in case of non-costly prosocial behavior that specifically benefits the other (i.e., in the *Disadvantageous prosocial* game), there was a decrease in the non-costly prosocial choices toward antagonists, neutral and anonymous peers. Thus, only for friends participants are willing to accept an unequal prosocial outcome. Such evidence for increasing as well as decreasing levels of prosocial behavior might help us better understand the previously reported contradictory findings on developmental patterns of prosocial behavior. Besides studies showing increasing levels of prosocial behavior (e.g., Eisenberg et al., 1991, 1995), there are findings suggesting a decline in prosocial behaviors from middle to late adolescence (e.g., Luengo Kanacri et al., 2013). Our findings suggest that future studies should better examine the role of interaction partners in displays of prosocial behavior to get a more nuanced idea on these developmental patterns.

The second set of analyses focused on trust and reciprocity in the Trust game. Contrary to expectations, trust- and reciprocity-related prosocial behavior showed no age related changes. That is to say, per interaction partner, participants of all ages showed similar levels of trust. Several prior developmental studies have demonstrated low levels of trust and reciprocity toward strangers in children and young adolescents, and that both trust and reciprocity behavior increase with age (Sutter and Kocher, 2007; van den Bos et al., 2010). The current findings add to this literature by showing that 9-year-olds can already display trust and reciprocity behavior when they are interacting with friends. Prior reports already indicated that interpersonal trust is an important aspect of friendship across childhood and adolescence (Bigelow and La Gaipa, 1975; Selman, 1980; Youniss, 1980). The reciprocal aspect of friendships increases in importance around elementary school and reciprocity remains to be the *deep structure* of friendships across the life-span (Hartup and Stevens, 1997).

A CLOSER LOOK AT THE ROLE OF INTERACTION PARTNERS IN PROSOCIAL BEHAVIOR

Young adolescents in the age group of 9- and 12-year-olds generally showed similar levels of prosocial behavior for all interaction partners. In contrast, 15- and 18-year-olds clearly differentiated in prosocial behavior depending on the interaction partner. When prosocial behavior was non-costly, 15- and 18-year-olds acted more prosocial toward friends and neutral peers than to disliked and anonymous ones; when it was costly, 18-year-olds further differentiated friends from neutral peers. Thus, the development of the differentiation of interaction partners in displays of costly prosocial behavior seems to be prolonged across adolescence; this might possibly be because prosocial behavior requires better control of self-outcome maximization.

Differentiation of interaction partners in displays of costly and non-costly prosocial behavior has been shown for 3.5- and 4.5-year-olds (Olson and Spelke, 2008; Moore, 2009). In light of these previous findings, it might be puzzling that 9- and 12-year-olds in our study did not differentiate at all between interaction partners. Our findings are further, however, in line with the findings of Buhrmester et al. (1992) where they show that 6- and 10-year-olds do not differentiate between friends and neutral peers in their sharing behavior, whereas 14-year-olds share more with friends than with neutral peers. The pattern of prosocial behavior of 15- and 18-year-olds in the current study, where we see the differentiation of friends from all other peers, fits well with the developmental role of friendships and their increasing importance across adolescence (Sullivan, 1953; Youniss, 1980).

The significant role of friendships across childhood and adolescence is further supported by the strong differentiation of friends from other peers in displays of trust and reciprocity. Participants of all ages showed highest levels of trust and reciprocity for friends. Oldest adolescents further differentiated between the other three peer groups, such that trust of anonymous and disliked peers were less often reciprocated than trust of neutral peers. It is noteworthy that even the youngest age groups differentiated between friends and other peers in their trust and reciprocity behavior, whereas this effect was lacking in the allocation games. It could be that trust and reciprocity develop initially within close relationships such as friendships, whereas fairness related prosocial behavior are more general forms of prosocial behavior that are not relationship-specific.

Interestingly, neither prosocial choices in the allocation games nor trust and reciprocity choices in the Trust game differed for disliked and anonymous peers in any of the age groups. It might be that within the current context both these groups were seen as an out-group and that adolescents differentiate mainly between in-group and out-group members of the peer group (Fehr et al., 2008). As Fehr et al. (2008) rightly indicate, prosocial behavior (particularly in the form of reciprocity) can be motivated by selfish impulses related to expectations of future benefits from interaction partners. In this respect, it could be that participants' lack of expectations to interact with disliked as well as with anonymous peers in the future might explain behavior in this context.

Taken together, across adolescence control of outcome-maximization and payoff comparisons are increasingly better incorporated into decision-making. These results are in line with our previous findings showing developmental patterns that are dependent on intentionality of unfair treatment (Güroğlu et al., 2009a,b, 2011; Overgaauw et al., 2012) and reputation based on

previous interactions (Will et al., 2013). These findings show that social context information is increasingly better incorporated into decision-making. Prior studies showed that, despite stable individual differences, prosocial behavior is difficult to predict over time (Eisenberg et al., 1999). Our findings suggest that prosocial behavior is increasingly sensitive to factors related to the social context in which interactions take place, which might explain weak consistency in prosocial behavior in prior studies.

ROLE OF PERSPECTIVE-TAKING IN PROSOCIAL BEHAVIOR

One of the questions that we addressed in this research was the role of perspective-taking skills as a possible mediator of age related differences in prosocial behavior. Indeed, we found that the age related increases in prosocial behavior toward friends was mediated by self-reported perspective taking in the *Non-costly prosocial* game. It has been shown that perspective-taking has a protracted developmental trajectory into late adolescence (Dumontheil et al., 2009). We provide further support for this developmental trajectory based on self-reported perspective-taking, and this pattern is linked to differences in prosocial behavior.

Considering that we found support for the mediating role of perspective-taking in age related increase in prosocial behavior only in one of the games examined here, caution must be taken in interpreting these results and their implications for generalization. Interestingly, only non-costly prosocial behavior was mediated by perspective-taking. Possibly, in the *Costly prosocial* game where prosocial behavior is costly, changes in other aspects of cognitive development, such as executive functioning and cognitive control, are more strongly related to costly prosocial behavior, where control of self-maximizing impulses play a role (Steinbeis et al., 2012; Luengo Kanacri et al., 2013). Although prosocial behavior in the *Disadvantageous prosocial* game was not costly, it can be considered as costly in terms of comparative interpersonal costs because it leads to a disadvantagous distribution of coins for the participant. In this sense, it could be that control of impulses also plays a relatively more important role than perspective-taking skills in this form of prosocial behavior. For future research it will be interesting to examine other interaction partners, such as parents, to better understand the aspects of social context that triggers perspective-taking and prosocial behavior. Previous studies also point out that perspective-taking skills play a significant role in both trust and reciprocity decision (Malhotra, 2004; van den Bos et al., 2010, 2011a,b). In the current study, due to practical considerations we could not employ similar study designs that would allow us to examine the role of perspective taking in trust and reciprocity. Future studies should aim to employ task manipulations that specifically address the role of perspective-taking in trust and reciprocity decision with different interaction partners.

CONCLUDING REMARKS

In the current study, we did not examine the role of gender in prosocial behavior in interactions with peers due to too small sample sizes per gender and age group. There is ample evidence on gender differences in both peer relationships and displays of prosocial behavior (Maccoby, 1986; Eisenberg et al., 1996). Across middle childhood and adolescence friendships are typically same-sex dyads, and friendships of girls are more often characterized by prosocial behavior, whereas friendships of boys more often involve displays of antisocial behavior (Güroğlu et al., 2007). Considering the relatively low prevalence of same-sex antipathy relationships (Güroğlu et al., 2009a,b), peer nominations were not restricted to same-sex nominations in the current study. Also, the small sample size within each age group did not allow us to examine gender effects. Future research should further examine the role of gender and gender combinations in peer interactions.

The experimental design of the allocation games in the current study ensured anonymity of all choices. This was done to restrict the possible role of social desirability in displays of prosocial behavior. A previous study examining sharing between friends and non-friends has shown that secret vs. public acts of sharing might differ (Buhrmester et al., 1992). Similarly, Leimgruber et al. (2012) provide evidence for strategic prosociality in 5-year-olds, where children behave more generously when the recipient is aware of the details of their actions. Considering that real-life social behavior usually takes place in the presence of others (peers, as well as parents and teachers), future studies should investigate how this aspect of context influences social decision-making across adolescence.

It is also important to note that 9-year-olds in our study do not differ in their frequency of prosocial behavior depending on the alternatives in each game. In other words, they do not differentiate between costly and non-costly prosocial behavior (see Supplementary Material). This is in contradiction with prior findings from similar and even younger age groups (Fehr et al., 2008; Blake and McAuliffe, 2011; Shaw and Olson, 2012), where decisions are shown to be affected by payoffs. Although the tasks were explained in detail to participants in small groups and all participants were given the chance to ask questions to ensure that everyone understood the task, it is possible that the youngest participants had trouble understanding the games. Future studies should include a comprehension check to assess whether children understand the payoff structure.

Here we employed a cross-sectional design to examine age differences in prosocial behavior. The current findings are highly informative for understanding developmental trajectories in prosocial behavior. Studies employing longitudinal designs are needed to reach conclusions regarding these developmental trajectories. Such longitudinal examinations will enable researchers to examine individual differences in peer relationship history (e.g., chronic rejection by peers or consistent popularity) and link these to cognitive changes (such as perspective-taking) and social behavior. However, longitudinal assessments of sociometric measures where complete school classes are tested using experimental designs as that employed here are challenging in terms of practical considerations. Future studies should focus on alternative ways of assessing prosocial behavior with real-life interaction partners that are feasible within longitudinal designs.

This study merges two important aspects of development: social decision-making and peer relationships. Our design is unique in the way it employs sociometric measures, a core method to assess peer relationships, and combines this with an experimental design using economic exchange games, which are highly efficient in examining social decision-making processes. The use of this experimental design employing allocation games tapping at different aspects of social decision-making further enabled us to examine prosocial behavior from different aspects, i.e., in terms of fairness, trust, and reciprocity considerations. The added value of this approach lies in its feasibility to examine social behavior toward different types of peers, which is not easily assessed using other methods such as questionnaire or observations of behavior. This approach is promising in understanding social exclusion in the peer context and the role of peer relationships in the treatment of bullies as well as victims (Güroğlu et al., 2013).

The differential patterns of behavior for interaction partners support the special role of friendships as forming the most significant developmental contexts across adolescence (Hartup, 1996), especially for prosocial behavior (Carlo et al., 1999). Converging evidence from all forms of behavior examined in this study is that adolescents treat friends differently than all other types of peers, and this special treatment is shaped throughout adolescence. In recent years, neuroscientific research has further highlighted the special and rewarding role of social interactions with friends (see e.g., Güroğlu et al., 2008; Braams et al., 2013). Future research needs to further pay attention to this context specificity of social behavior, and examine its links with the developing social brain.

ACKNOWLEDGMENTS

The authors would like to thank Landa Endlich, Brenda Riegman, Laura Stevens, and Marjolijn van Woudenberg for their help with the data collection. This research was supported by the Netherlands Organization for Scientific Research (NWO) Grants 056-34-010 to Eveline A. Crone and 451-10-021 to Berna Güroğlu.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00291/abstract

REFERENCES

Amato, P. R. (1990). Personality and social network involvement as predictors of helping behavior in everyday life. *Soc. Psychol. Q.* 53, 31–43. doi: 10.2307/2786867

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity and social history. *Games Econ. Behav.* 10, 122–142. doi: 10.1006/game.1995.1027

Berndt, T. J. (1981). Age changes and changes over time in prosocial intentions and behavior between friends. *Dev. Psychol.* 17, 408–416. doi: 10.1037/0012-1649.17.4.408

Berndt, T. J. (1985). Prosocial behavior between friends in middle childhood and early adolescence. *J. Early Adolesc.* 5, 307–317. doi: 10.1177/0272431685053005

Bigelow, B. J., and La Gaipa, J. J. (1975). Children's written descriptions of friendship: a multidimensional analysis. *Dev. Psychol.* 11, 857–858. doi: 10.1037/0012-1649.11.6.857

Blake, P. R., and McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition* 120, 215–224. doi: 10.1016/j.cognition.2011.04.006

Braams, B., Güroğlu, B., de Water, E., Meuwese, R., Koolschijn, P. C., Peper, J., et al. (2013). Reward-related neural responses are dependent on the beneficiary. *Soc. Affect. Cogn. Neurosci.* doi: 10.1093/scan/nst077. [Epub ahead of print].

Brown, B. B. (2004). "Adolescents' relationships with peers," in *Handbook of Adolescent Psychology,* eds R. M. Lerner and L. Steinberg (New Jersey: John Wiley), 363–394.

Buhrmester, D., Goldfarb, J., and Cantrell, D. (1992). Self-presentation when sharing with friends and nonfriends. *J. Early Adolesc.* 12, 61–79. doi: 10.1177/0272431692012001004

Burnham, T., McCabe, K., and Smith, V. L. (2000). Friend-or-foe intentionality priming in an extensive form trust game. *J. Econ. Behav. Organ.* 43, 57–73. doi: 10.1016/S0167-2681(00)00108-6

Card, N. A., and Hodges, E. V. E. (2007). Victimization within mutually antipathetic relationships. *Soc. Dev.* 16, 479–496. doi: 10.1111/j.1467-9507.2007.00394.x

Carlo, G., Fabes, R. A., Laible, D., and Kupanoff, K. (1999). Early adolescence and prosocial/moral behavior II: the role of social and contextual influences. *J. Early Adolesc.* 19, 133–147. doi: 10.1177/0272431699019002001

Carlo, G., and Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescents. *J. Youth Adolesc.* 31, 31–44. doi: 10.1023/A:1014033032440

Carpenter, P. A., Just, M. A., and Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychol. Rev.* 97, 404–431. doi: 10.1037/0033-295X.97.3.404

Costin, S. E., and Jones, D. C. (1992). Friendship as a facilitator of emotional responsiveness and prosocial interventions among young children. *Dev. Psychol.* 28, 941–947. doi: 10.1037/0012-1649.28.5.941

Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *J. Pers. Soc. Psychol.* 44, 113–126. doi: 10.1037/0022-3514.44.1.113

Delaney, H. D., and Maxwell, S. E. (1981). On using analysis of covariance in repeated measures designs. *Multivariate Behav. Res.* 16, 105–123. doi: 10.1207/s15327906mbr1601_6

Delgado, M. R., Frank, R. H., and Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618. doi: 10.1038/nn1575

Dumontheil, I., Apperly, I. A., and Blakemore, S.-J. (2009). Online usage of theory of mind continues to develop in late adolescence. *Dev. Sci.* 13, 331–338. doi: 10.1111/j.1467-7687.2009.00888.x

Dunham, Y., Baron, A. S., and Carey, S. (2011). Consequences of 'minimal' group affiliations in children. *Child Dev.* 82, 793–811. doi: 10.1111/j.1467-8624.2011.01577.x

Eisenberg, N., Carlo, G., Murphy, B. C., and van Court, P. (1995). Prosocial development in late adolescence: a longitudinal study. *Child Dev.* 66, 1179–1197. doi: 10.2307/1131806

Eisenberg, N., Fabes, R. A., Karbon, M., Murphy, B. C., Woskinski, M., Polazzi, L., et al. (1996). The relations of children's dispositional prosocial behavior to emotionality, regulation, and social functioning. *Child Dev.* 67, 974–992. doi: 10.2307/1131874

Eisenberg, N., Fabes, R. A., and Spinrad, T. L. (2006). "Prosocial behavior," in *Handbook of Child Psychology:* Vol. 3. Social, emotional, and personality development, 6th Edn., W. Damon and R. M. Lerner (Series ed) and N. Eisenberg, (Vol. ed), (New York, NY: Wiley), 646–718.

Eisenberg, N., Guthrie, I. K., Murphy, B. C., Shepard, S. A., Cumberland, A., and Carlo, G. (1999). Consistency and development of prosocial dispositions: a longitudinal study. *Child Dev.* 70, 1360–1372. doi: 10.1111/1467-8624.00100

Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behavior. *Psychol. Bull.* 101, 91–119. doi: 10.1037/0033-2909.101.1.91

Eisenberg, N., Miller, P. A., Shell, R., McNalley, S., and Shea, C. (1991). Prosocial development in adolescence: a longitudinal study. *Dev. Psychol.* 27, 849–857. doi: 10.1037/0012-1649.27.5.849

Fehr, E., Bernhard, H., and Rockenbach, B. (2008). Egalitarianism in young children. *Nature* 454, 1079–1084. doi: 10.1038/nature07155

Fehr, E., Fischbacher, U., and Gächter, S. (2002). Strong reciprocity, human cooperation and the enforcement of social norms. *Hum. Nat.* 13, 1–25. doi: 10.1007/s12110-002-1012-7

Furman, W., and Bierman, K. L. (1984). Children's conceptions of friendship: a multidimentional study of developmental changes. *Dev. Psychol.* 20, 925–931. doi: 10.1037/0012-1649.20.5.925

Gummerum, M., Keller, M., Takezawa, M., and Mata, J. (2008). To give or not to give: children's and adolescents' sharing and moral negotiations in economic decision situations. *Child Dev.* 79, 562–576. doi: 10.1111/j.1467-8624.2008.01143.x

Güroğlu, B., Haselager, G. J. T., van Lieshout, C. F. M., Takashima, A., Rijpkema, M., and Fernández, G. (2008). Why are friends special? Implementing a social interaction simulation task to probe the neural correlates of friendship. *Neuroimage* 39, 903–910. doi: 10.1016/j.neuroimage.2007.09.007

Güroğlu, B., Haselager, G., van Lieshout, C. F. M., and Scholte, R. H. J. (2009a). Antagonists in antipathy relationships: a person-oriented approach. *J. Res. Adolesc.* 19, 35–46. doi: 10.1111/j.1532-7795.2009.00580.x

Güroğlu, B., van den Bos, W., and Crone, E. A. (2009b). Fairness considerations: Increasing understanding of intentionality in adolescence. *J. Exp. Child Psychol.* 104, 398–409. doi: 10.1016/j.jecp.2009.07.002

Güroğlu, B., van den Bos, W., van Dijk, E., Rombouts, S. A. R. B., and Crone, E. A. (2011). Dissociable brain networks involved in development of fairness considerations: understanding intentionality behind unfairness. *Neuroimage* 57, 634–641. doi: 10.1016/j.neuroimage.2011.04.032

Güroğlu, B., van Lieshout, C. F. M., Haselager, G. J. T., and Scholte, R. H. J. (2007). Similarity and complementarity of behavioral profiles of friendship types and types of friends: friendships and psychosocial adjustment. *J. Res. Adolesc.* 17, 357–386. doi: 10.1111/j.1532-7795.2007.00526.x

Güroğlu, B., Will, G.-J., and Klapwijk, E. T. (2013). Some bullies are more equal than others: peer relationships modulate altruistic punishment of bullies after observing ostracism. *Int. J. Dev. Sci.* 7, 13–23. doi: 10.3233/DEV-1312117

Hartup, W. (1996). The company they keep: friendships and their developmental significance. *Child Dev.* 67, 1–13. doi: 10.2307/1131681

Hartup, W. W., and Stevens, N. (1997). Friendships and adaptation in the life course. *Psychol. Bull.* 121, 355–370. doi: 10.1037/0033-2909.121.3.355

Hoffman, M. L. (1975). Developmentl synthesis of affect and cognition and its implications for altruistic motivation. *Dev. Psychol.* 11, 607–622. doi: 10.1037/0012-1649.11.5.607

Hubbard, J., Dodge, K., Cillessen, A., Coie, J., and Schwartz, D. (2001). The dyadic nature of social information processing in boys' reactive and proactive aggression. *J. Pers. Soc. Psychol.* 80, 268–280. doi: 10.1037/0022-3514.80.2.268

Iannotti, R. J. (1985). Naturalistic and structured assessments of prosocial behavior in preschool children: the influence of empathy and perspective taking. *Dev. Psychol.* 21, 46–55. doi: 10.1037/0012-1649.21.1.46

Lahno, B. (1995). Trust, reputation, and exit in exchange relationships. *J. Confl. Resolut.* 39, 495–510. doi: 10.1177/0022002795039003005

Larson, A. (1992). Network dyads in entrepreneurial settings: a study of the governance of exchange relationships. *Adm. Sci. Q.* 37, 76–104. doi: 10.2307/2393534

Leimgruber, K. L., Shaw, A., Santos, L. R., and Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLoS ONE* 7:e48292. doi: 10.1371/journal.pone.0048292

Luengo Kanacri, B. P., Pastorelli, C., Eisenberg, N., Zuffianò, A., and Caprara, G. V. (2013), The development of prosociality from adolescence to early adulthood: the role of effortful control. *J. Pers.* 81, 302–312. doi: 10.1111/jopy.12001

Maccoby, E. E. (1986). "Social groupings in childhood: their relationship to prosocial and antisocial behavior in boys and girls," in *Development of Antisocial and Prosocial Behaviour: Research, Theories and Issues*, eds D. Olweus, J. Block, and M. Radke-Yarrow (New York, NY: Academic Press Inc), 263–284.

Malhotra, D. (2004). Trust and reciprocity decisions: the differing perspectives of trustors and trusted parties. *Organ. Behav. Hum. Decis. Process.* 94, 61–73. doi: 10.1016/j.obhdp.2004.03.001

Moore, C. (2009). Fairness in children's resource allocation depends on the recipient. *Psychol. Sci.* 20, 944–949. doi: 10.1111/j.1467-9280.2009.02378.x

Moore, C., Barresi, J., and Thompson, C. (1998). The cognitive basis of future-oriented prosocial behavior. *Soc. Dev.* 7, 198–218. doi: 10.1111/1467-9507.00062

Nelson, D. A., and Crick, N. R. (1999). Rose-colored glasses: examining the social information-processing of prosocial young adolescents. *J. Early Adolesc.* 19, 17–38. doi: 10.1177/0272431699019001002

Nummenmaa, L., Peets, K., and Salmivalli, C. (2008). Automatic activation of adolescents' peer-relational schemas: evidence from priming with facial identity. *Child Dev.* 79, 1659–1675. doi: 10.1111/j.1467-8624.2008.01217.x

Olson, K. R., and Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition* 108, 222–231. doi: 10.1016/j.cognition.2007.12.003

Overgaauw, S., Güroğlu, B., and Crone, E. A. (2012). Fairness considerations when I know more than you do: developmental comparisons. *Front. Dev. Psychol.* 3:424. doi: 10.3389/fpsyg.2012.00424

Peets, K., Hodges, E. V. E., Kikas, E., and Salmivalli, C. (2007). Hostile attributions and behavioral strategies in children: does relationship type matter? *Dev. Psychol.* 43, 889–900. doi: 10.1037/0012-1649.43.4.889

Preacher, K. J., and Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav. Res. Methods Instrum. Comput.* 36, 717–731. doi: 10.3758/BF03206553

Radke, S., Güroğlu, B., and de Bruijn, E. R. A. (2012). There's something about a fair split: intentionality moderates context-based fairness considerations in social decision-making. *PLoS ONE* 7:e31491. doi: 10.1371/journal.pone.0031491

Ray, G. E., and Cohen, R. (1997). Children's evaluations of provocation between peers. *Aggress. Behav.* 23, 417–431.

Rilling, J. K., and Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annu. Rev. Psychol.* 62, 23–48. doi: 10.1146/annurev.psych.121208.131647

Sebastián-Enesco, C., Hernández-Lloreda, M. V., and Colmenares, F. (2013). Two and a half-year-old children are prosocial even when their partners are not. *J. Exp. Child Psychol.* 116, 186–198. doi: 10.1016/j.jecp.2013.05.007

Selman, R. L. (1980). *The Growth of Interpersonal Understanding*. New York, NY: Academic Press.

Shaw, A., and Olson, K. R. (2012). Children discard a resource to avoid inequity. *J. Exp. Child Psychol.* 141, 382–395. doi: 10.1037/a0025907

Smith, C. E., Blake, P. R., and Harris, P. L. (2013). I should but I won't: why young children endorse norms of fair sharing but do not follow them. *PLoS ONE* 8:e59510. doi: 10.1371/journal.pone.0059510

Steinbeis, N., Bernhardt, B. C., and Singer, T. (2012). Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron* 73, 1040–1051. doi: 10.1016/j.neuron.2011.12.027

Steinbeis, N., and Singer, T. (2013). The effects of social comparison on social emotions and behavior during childhood: the ontogeny of envy and Schadenfreude predicts developmental changes in equity-related decisions. *J. Exp. Child Psychol.* 115, 198–209. doi: 10.1016/j.jecp.2012.11.009

Sullivan, H. S. (1953). *The Interpersonal Theory of Psychiatry*. New York, NY: W. W. Norton.

Sutter, M. (2007). Outcomes versus intentions: on the nature of fair behavior and its development with age. *J. Econ. Psychol.* 28, 69–78. doi: 10.1016/j.joep.2006.09.001

Sutter, M., and Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games Econ. Behav.* 59, 364–382. doi: 10.1016/j.geb.2006.07.006

van den Bos, W., van Dijk, E., and Crone, E. A. (2011a). Learning whom to trust in repeated social interactions: a developmental perspective. *Group Processes Intergroup Relat.* 15, 243–256. doi: 10.1177/1368430211418698

van den Bos, W., van Dijk, E., Westenberg, P. M., Rombouts, S. A. R. B. and Crone, E. A. (2011b). Changing brains, changing perspectives: the neurocognitive development of reciprocity. *Psychol. Sci.* 22, 60–70. doi: 10.1177/0956797610391102

van den Bos, W., Westenberg, M., van Dijk, E., and Crone, E. A. (2010). Development of trust and reciprocity in adolescence. *Cogn. Dev.* 25, 90–102. doi: 10.1016/j.cogdev.2009.07.004

van Leijenhorst, L., Westenberg, M., and Crone, E. A. (2008). A developmental study of risky decisions on the cake gambling task: age and gender analyses of probability estimation and reward evaluation. *Dev. Neuropsychol.* 33, 179–196. doi: 10.1080/87565640701884287

Warneken, F., Lohse, K., Melis, A. P., and Tomasello, M. (2011). Young children share the spoils after collaboration. *Psychol. Sci.* 22, 267–273. doi: 10.1177/0956797610395392

Will, G. J., Crone, E. A., van den Bos, W., and Güroğlu, B. (2013). Acting on observed social exclusion: developmental perspectives on punishment

of excluders and compensation of victims. *Dev. Psychol.* doi: 10.1037/a0032299

Youniss, J. (1980). *Parents and Peers in Social Development: A Sullivan-Piaget Perspective*. Chicago, IL: University of Chicago Press.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 November 2013; accepted: 20 March 2014; published online: 11 April 2014.

*Citation: Güroğlu B, van den Bos W and Crone EA (2014) Sharing and giving across adolescence: an experimental study examining the development of prosocial behavior. Front. Psychol. **5**:291. doi: 10.3389/fpsyg.2014.00291*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Güroğlu, van den Bos and Crone. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 29 April 2014
doi: 10.3389/fpsyg.2014.00388

The early development of executive function and its relation to social interaction: a brief review

Yusuke Moriguchi [1,2] *

[1] Department of School Education, Joetsu University of Education, Joetsu, Japan
[2] Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, Tokyo, Japan

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Li Qu, Nanyang Technological University, Singapore
Jean-Claude Croizet, University of Poitiers, France

***Correspondence:**
Yusuke Moriguchi, Department of School Education, Joetsu University of Education, 1 Yamayashiki-machi, Joetsu 943-8512, Japan; Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, Tokyo, Japan
e-mail: moriguchi@juen.ac.jp

Executive function (EF) refers to the ability to execute appropriate actions and to inhibit inappropriate actions for the attainment of a specific goal. Research has shown that this ability develops rapidly during the preschool years. Recently, it has been proposed that research on EF should consider the importance of social interaction. In this article, recent evidence regarding the early development of EF and its relation to social interaction has been reviewed. Research consistently showed that social interaction can influence EF skills in young children. However, the development of EF may facilitate the cognitive skills that are important for social interaction. Taken together, there might be functional dependency between the development of EF and social interaction.

Keywords: executive function, social interaction, preschool children, theory of mind, cognitive development

INTRODUCTION

Executive function (EF) is a complex cognitive control responsible for making adaptive changes in physical and social environments. It enables us to execute appropriate actions, and to inhibit inappropriate actions, to attain a goal (Dempster, 1992). Extensive evidence suggests that EF develops rapidly in the preschool years, with adult-level performance being achieved during adolescence (Anderson, 2002; Zelazo et al., 2003). The development of EF is supported by the maturation of the prefrontal cortex in preschool children as well as school-aged children (Diamond, 2002; Durston et al., 2006; Moriguchi and Hiraki, 2009).

One important issue in EF research is its structure. Adult research has shown that EF is not unitary. It consists of some sub-components, such as inhibition, shifting, and updating (working memory; Miyake et al., 2000; Miyake and Friedman, 2012). Although their studies focused on healthy adult populations, studies for elderly people and school-aged children also confirmed the Miyake et al. (2000) model (Lehto et al., 2003; Huizinga et al., 2006). However, in preschool children, a single-factor model (general EF) may be more appropriate (Wiebe et al., 2008). Alternatively, in younger children, the model of "conflict" EF and "delay" EF may be useful. The former refers to inhibiting a prepotent response while activating conflicting novel responses. The conflict EF is indexed by Stroop-like tasks or rule switching tasks. The latter refers to simply inhibiting responding, which is indexed by a delay of gratification (Carlson and Moses, 2001).

In the previous studies, problem-solving tasks, such as the Dimensional Change Card Sort (DCCS) task or the Day–Night Stroop task (Gerstadt et al., 1994; Zelazo et al., 1996; Kirkham et al., 2003) were extensively used. In the DCCS task, children are asked to sort cards that have two dimensions such as color and shape (e.g., red boats, blue rabbits). There are two phases in the task. In the first phase, children are asked to sort cards according to one dimension (e.g., color), for five or six trials. In the second phase, they are asked to sort the cards according to the other dimension (e.g., shape), for five or six trials. Research has shown that 3-year-olds tend to perseverate on the first dimension whereas older children do not show such perseveration. The researchers argued how and why young children made perseverative errors in the DCCS (Kirkham et al., 2003; Zelazo et al., 2003; Kloo and Perner, 2005).

Moreover, recently, the neural basis of EF in young children has been extensively examined (Moriguchi and Hiraki, 2011, 2014; Espinet et al., 2012; Buss et al., 2014). Indeed, research using near-infrared spectroscopy has shown that the performances of the DCCS tasks were significantly correlated with activations in the lateral prefrontal areas (Moriguchi and Hiraki, 2009). Further, the amplitude of N2 components measured by event-related potentials differed between children who passed and failed the DCCS tasks (Espinet et al., 2012). Previous research consistently showed that activations in the prefrontal cortex are important for successful performances in EF tasks.

SOCIAL INTERACTION INFLUENCES EF IN LATER DEVELOPMENT

However, until recently, how EF develops in social interaction has been largely neglected. This is in spite of the fact that several theorists proposed that humans are in nature social and develop through social interaction (Vygotsky, 1978; Tomasello, 2009). It has been proposed that higher mental functions, such as self-regulation, develop within the context of interpersonal activity (Vygotsky, 1978). In this view, interpersonal interaction may facilitate internalizing some views of another person's perspective on reality, which improves the development of higher mental

functions. This process is clearly manifested in a parent-child relationship. Parents provide children with alternative perspectives regarding how to deal with a given problem that can be internalized by a child. When children define an inappropriate goal, parents may direct children's attention to an appropriate goal. The external dialog between parent and child may go on to be internalized by the child in later development.

Lewis and Carpendale (2009) proposed that research on EF should consider the roles of social interaction. Indeed, there is evidence that social interaction between parents and children influences the development of EF. According to Roskam et al. (2014), there might be two possible dimensions of how interaction between parents and children influences EF in later development: supportive parenting and negative control. The former includes scaffolding, acceptance, and autonomy, which may facilitate children's development of EF. For example, Landry et al. (2002) showed that maternal verbal scaffolding affected EF skills such as search retrieval, mediated by children's verbal and non-verbal problem-solving skills. Bernier et al. (2010) examined whether maternal sensitivity, mind-mindedness and scaffolding at 12 to 15 months of age predicted children's EF skills, such as working memory and set shifting. Sensitivity refers to the tendency to read the child's needs and respond sensitively and appropriately (Ainsworth et al., 1978). Mind-mindedness refers to maternal appropriate use of mental language to their infants (Meins et al., 2003). The results revealed that scaffolding was the strongest predictor of the development of EF. Moreover, Hughes and Ensor (2009) reported that maternal scaffolding as well as other factors, such as imitative learning, plays an important role in the development of EF.

The other dimension about the relationship between parenting and EF in later development may be negative control. Negative control refers to parenting through coercion and punishment. It has been repeatedly shown that such parenting may lead to children's negative behaviors in later development (Gershoff, 2002). In terms of EF, Roskam et al. (2014) reported that negative control parenting may have negative impact on children's EF skills in later development. Specifically, the longitudinal research has shown that changes in negative parenting (e.g., frequent use of punishment) induced negative changes in inhibitory skills. Similarly, Blair et al. (2011) reported that positive (e.g., sensitivity) and negative (e.g., intrusiveness) parenting during infancy influenced EF and IQ in later development. The effect of negative parenting may be due to that such parenting may fail to provide children with the opportunity to control their actions.

Despite insufficient evidence, the previous studies suggest that social interaction, specifically interaction with parents, influenced the development EF skills in young children. Although the previous studies have examined the effect of parenting on children's EF, interaction with peers can be also important. Indeed, some researchers proposed that collaborative learning can enhance cognitive development, such as traditional Piagetian tasks (Doise and Mugny, 1984). In collaborative learning, each individual have different opinion in a given tasks, which induces social conflict. However, such situations may provide children with the impetus to have different perspectives, which may lead to improvement of the performances. According to Qu (2011), there are several reasons why collaboration with another person can facilitate children's executive control. For example, children may be aware a goal of the task and can have another person's perspectives thorough collaboration with another person, which lead to more efficient executive control. Future research should focus on the role of peer interaction in the development of EF.

SOCIAL LEARNING AND EF

The research above showed that parenting may play an important role in EF skills in later development. Such previous research clarified the long-term relationship between EF and social interaction, but it is still unclear how social interaction directly influenced children's behaviors that require executive control. Thus, I introduce the experimental evidence in the context of social and imitative learning. It is well known that children "overimitate" another person's behaviors (Horner and Whiten, 2005; Lyons et al., 2007; McGuigan et al., 2007). Overimitation refers to children's tendency to reproduce an adult's obviously irrelevant actions. For example, in a tool-use task, chimpanzees and 3- to 4-year-old children were asked to observe an experimenter using a tool to obtain a reward from a complex-structured box (Horner and Whiten, 2005). In the demonstration, some actions were causally relevant to obtain a reward, and other actions were causally irrelevant. When they performed the actions, chimpanzees only reproduced relevant actions whereas human children reproduced both relevant and irrelevant actions. Research on overimitation suggests that social and imitative learning may be so powerful that children may fail to control their behaviors after such learning.

Moriguchi et al. (2007) examined whether children's executive control might be influenced by learning from another person's actions. They modified the DCCS. In the modified social DCCS task, during the first phases, instead of sorting the cards by themselves, preschoolers watched an adult model sorting the cards according to one dimension (e.g., shape). During the second phases, children were asked to sort according to a different dimension (e.g., color). The results showed that most 3-year-olds perseverated sorting according to the observed dimension, as in the standard DCCS task. Thus, 3-year-old children used the observed rules even though they were asked to use the different rules. On the other hand, more than half of the 4-year-old children and most of the 5-year-old children did not use the observed rules, and sorted the cards according to the instructed, second rules.

Interestingly, children's behaviors were significantly affected by a model's mental states (Moriguchi et al., 2007). For example, children were more likely to use the observed rules when they observed a model who was confident with the rule she used than when they observed a model who lacked confidence with the rules. Moreover, the performance on the modified DCCS tasks is significantly correlated with performance on the standard DCCS tasks (Moriguchi and Itakura, 2008).

Children's executive control process could be affected by a human's actions, but not a robot's actions. Moriguchi et al. (2010) showed that children who observed a robot sorting according to one dimension had no difficulty in sorting the cards according to a different dimension. Moriguchi et al. (2010) reported that

the effects of demonstration by an android (a robot with human appearance) were stronger than those by a robot, but weaker than those by a human model. The authors explain the results in terms of a sociocognitive perspective that children perseverate on the human model's rule because they mentally simulate the model's actions while watching. On the other hand, the children's actions were not affected by the robot's actions because the robot did not induce young children's simulative processes.

There might be some cultural differences in performance on the social DCCS. Moriguchi et al. (2012) gave 3- and 4-year-olds in Canada and Japan the standard version of the DCCS and the social version of the DCCS. Results indicated that Canadian children displayed the perseverative behaviors in the social DCCS, but their effects were relatively weaker than those of Japanese children. On the other hand, performance on the standard DCCS was similar between the two countries. Although the general developmental trajectory may be common in two cultures, the results can be interpreted in terms of cultural psychology theories. People in Western cultures tend to have a more "independent" view of the self, whereas people in Asian cultures are likely to have a more "interdependent" view (Markus and Kitayama, 1991). In interdependent cultures, people may recognize that their behaviors can be affected by others' behaviors. On the other hand, in independent cultures, people may tend to believe that their behaviors are independent from others'. Thus, Canadian children may be more likely to separate themselves from another person than Japanese children.

The effects of social interaction on EF were reported using a Stroop-like Black/White task (Moriguchi, 2012). In this task, children were asked to respond to a pair of pictures: in the black/white task, for example, children had to respond "black" when shown a white card, and respond "white" when shown a black card. Children were told to suppress the tendency to respond according to what color the card was and instead activate a conflicting response (Simpson and Riggs, 2005). This study compared the standard condition to an interference condition. In the interference condition, children observed incorrect demonstrations, where the demonstrator responded with "black" to a black card, and "white" to a white card; they were then given the black/white task. The results revealed that children in the interference condition performed worse than those in the neutral condition.

Research suggests that interaction with a human, but not with non-human agents, can affect children's EF skills. In addition, culture may influence the relationship between EF and social interaction. Taken together with the evidence of parenting, social interaction may have a strong impact on children's executive control.

EF INFLUENCES SOCIAL INTERACTION

The research reported above suggests that social interaction can affect children's EF skills. Conversely, there is accumulating evidence that EF skills may facilitate the development of social interaction. The most well-known case is that EF was significantly correlated with theory of mind (ToM; Frye et al., 1995; Hughes, 1998; Carlson and Moses, 2001; Sabbagh et al., 2006; Benson et al., 2012). ToM refers to the ability of children to be aware that they or other individuals can have mental states, such as false beliefs. Extensive research indicates that representative false belief understanding improves markedly during the preschool years (Wellman et al., 2001).

Given the existing evidence regarding the relationship between EF and ToM, theorists argued how EF was related to ToM. First, the development of false belief understanding may contribute to improvement in children's EF (Perner et al., 2002). According to the view, metarepresentational understanding underlying ToM provides the foundation for the development of EF skills. Indeed, Kloo and Perner (2003) reported that training children on ToM tasks leads to improvement in their performance on the DCCS task. Nevertheless, DCCS training also improves children's performance on ToM tasks. Moreover, longitudinal research has shown that early EF skills (i.e., 2 years of age) predict later ToM abilities (i.e., at 3 years of age) rather than the reverse (Carlson et al., 2004). This evidence may challenge the view that ToM improves EF skills.

Many researchers have speculated that conflict EF is fundamental for the development of false belief understanding. There are mainly two explanations in this view. One explanation is the expression view, by which children fail to perform false belief tasks because children did not have EF skills to deal with the task demands in the false belief tasks. In this view, children do have an understanding of another person's false belief, but appeared to lack understanding due to their poor executive skills. The other explanation is the emergence view. On this view, EF may be necessary for the emergence of children's false belief understanding. The recent evidence favors the latter view (Benson et al., 2012; Devine and Hughes, 2014). For example, children's EF skills are correlated with performance on ToM tasks with fewer executive demands (Henning et al., 2011).

In relation to this point, EF in young children might be correlated with their lying behaviors. Lying involves a false statement with the intention to deceive the recipient while considering the recipient's psychological state (e.g., knowledge). Talwar and Lee (2008) administered a peeking task, EF tasks, and ToM tasks to 3- to 8-year-old children, and examined the relationship between them. In the peeking task, after the children and an experimenter played a game, the experimenter left the room. The children were told not to peek at a toy while the experimenter was out of the room. After the experimenter returned to the room, children were asked whether they had peeked at the toy. There were two questions. The first question was whether they had peeked at the toy, and the second question was what the toy was. The assumption on the second questions was that if children had not peeked at it, they would not know what the toy was. The results revealed that children's lying in respond to the first question about peeking was significantly correlated with their EF skills measured by the Day–Night Stroop task and the scores of false belief tasks. The lying in response to the second question was also correlated with the scores of the Day–Night Stroop task. That is, children who developed more EF skills tended to lie more.

In terms of the relationship between EF and communicative behaviors, Moriguchi et al. (2008) examined the correlation between the performance on the DCCS tasks and a yes bias in

preschool children. The yes bias is children's tendency to answer "yes" when they are posed yes/no questions. The bias occurs in spite of knowing that the correct answer is "no" (Fritzley and Lee, 2003; Okanda and Itakura, 2007). Okanda and Itakura (2007) suggested that affirmation including a yes response could be a dominant response and children would not able to inhibit the response. Thus, it was possible that having inhibitory skills help a child to avoid saying the first thing that comes to his or her mind when asked a question by an interviewer. Moriguchi et al. (2008) found that better inhibitory control ability was significantly related to a weaker yes bias even after controlling for age and verbal ability.

Other research showed that EF might be correlated with the development of moral behaviors. Kochanska et al. (1996, 1997) examined the relationship between effortful control (more temperamental aspects of EF) and moral-related behaviors in young children. For example, children's effortful control assessed by tasks such as delay of gratification were significantly correlated with children's internalizations of mothers' prohibitions of refraining from attractive activities (Kochanska et al., 1996). In addition, Kochanska et al. (1997) reported that children's effortful control at toddler and preschool age longitudinally contributed to conscience development at an early school age. The conscience development was measured by sustaining mundane activities and suppressing desired behaviors. Moreover, children's effortful control was significantly negatively correlated with antisocial responses on hypothetical moral dilemma tests. Further, children's views of themselves on moral dimensions were significantly correlated with effortful control.

Taken together, the previous research showed that EF skills are significantly correlated with false belief understanding, lying behaviors, responses in questioning, and the internalization of rules in young children. The causal relationship between these variables is still unclear. It is possible that the development of EF contributes to socio-communicative behaviors, or vice versa, and therefore future research should address the causal relationship. Nevertheless, the previous results suggest that the development of EF is closely related to the development of cognitive skills important for social interaction.

CONCLUSION AND FUTURE DIRECTION

In sum, previous studies showed that social interaction may facilitate the development of EF. Specifically, maternal scaffolding may be a strong predictor of the development of EF skills. In addition, interaction with a person, not a non-human agent, can be important for children's EF skills. Conversely, children's EF skills may also facilitate the development of social interaction skills, such as ToM, communicative behaviors, and moral skills. Taken together, the accumulated evidence suggests functional dependency between the development of EF skills and social interaction.

Future research should utilize social interaction to intervene with children who have lower EF skills. Recently, it has been repeatedly shown that self-control abilities, including EF skills, in young children predict school success and socioeconomic status during adulthood (Blair and Razza, 2007; Moffitt et al., 2011). Thus, several training programs have been proposed to facilitate children's EF skills (Lillard and Else-Quest, 2006; Thorell et al., 2009; Diamond and Lee, 2011). Some trainings use computer-based programs, and others used school curricula. Given the present review, we suggest that intervention programs that include social interaction can be more useful than those that do not.

The other possible direction for future research is to examine the neural basis of EF skills and its relation to social interaction. The development of EF skills is related to the activations in the prefrontal cortex (Espinet et al., 2012; Moriguchi and Hiraki, 2013). However, it is still unclear which factors affect the development of the prefrontal activations. Given the evidence reported here, maternal scaffolding can affect the development of the prefrontal activations in young children. Moreover, there is an argument regarding whether EF skills share brain regions with social interaction skills, such as ToM (Perner and Lang, 2000). There may be some commonalities in the brain regions, although core brain regions in EF may be different from those in ToM (Miller and Cohen, 2001; Saxe et al., 2006; Kalbe et al., 2010). The previous research was mostly based on adult brain imaging research or neuropsychological evidence, and there was little neuroimaging research in young children. Given that some of the brain regions may begin with broad functionality, and then be specialized to a given stimuli and task (Johnson, 2011), it is possible that the neural basis of EF shares the neural basis of ToM in young children. Thus, future studies should address these possibilities.

ACKNOWLEDGMENTS

This study was supported by a grant from the JST PRESTO program. The research was also funded by a grant from JSPS KAKENHI Grant Number 22700269, 24650133.

REFERENCES

Ainsworth, M. D. S., Blehar, M. C., Waters, E., and Wall, S. (1978). *Patterns of Attachment: A Psychological Study of The Strange Situation.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724

Benson, J. E., Sabbagh, M. A., Carlson, S. M., and Zelazo, P. D. (2012). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Dev. Psychol.* 49, 1615–1627. doi: 10.1037/a0031056

Bernier, A., Carlson, S. M., and Whipple, N. (2010). From external regulation to self-regulation: early parenting precursors of young children's executive functioning. *Child Dev.* 81, 326–339. doi: 10.1111/j.1467-8624.2009.01397.x

Blair, C., Granger, D. A., Willoughby, M., Mills-Koonce, R., Cox, M., Greenberg, M. T., et al. (2011). Salivary cortisol mediates effects of poverty and parenting on executive functions in early childhood. *Child Dev.* 82, 1970–1984. doi: 10.1111/j.1467-8624.2011.01643.x

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Buss, A. T., Fox, N., Boas, D. A., and Spencer, J. P. (2014). Probing the early development of visual working memory capacity with functional near-infrared spectroscopy. *Neuroimage* 85(Pt 1), 314–325. doi: 10.1016/j.neuroimage.2013.05.034

Carlson, S. M., Mandell, D. J., and Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Dev. Psychol.* 40, 1105–1122. doi: 10.1037/0012-1649.40.6.1105

Carlson, S. M., and Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism – toward a unified theory of cognitive-development and aging. *Dev. Rev.* 12, 45–75. doi: 10.1016/0273-2297(92)90003-K

Devine, R. T., and Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: a meta-analysis. *Child Dev.* doi: 10.1111/cdev.12237 [Epub ahead of print].

Diamond, A. (2002). "Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry," in *Principles of Frontal Lobe Function,* eds D. T. Stuss and R. T. Knight (New York: Oxford University Press), 466–503.

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529

Doise, W., and Mugny, G. (1984). *The Social Development of The Intellect.* Oxford: Pergamon Press.

Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J. A., et al. (2006). A shift from diffuse to focal cortical activity with development. *Dev. Sci.* 9, 1–8. doi: 10.1111/j.1467-7687.2005.00454.x

Espinet, S. D., Anderson, J. E., and Zelazo, P. D. (2012). N2 amplitude as a neural marker of executive function in young children: an ERP study of children who switch versus perseverate on the dimensional change card sort. *Dev. Cogn. Neurosci.* 2(Suppl. 1), S49–S58. doi: 10.1016/j.dcn.2011.12.002

Fritzley, V. H., and Lee, K. (2003). Do young children always say yes to yes-no questions? A metadevelopmental study of the affirmation bias. *Child Dev.* 74, 1297–1313. doi: 10.1111/1467-8624.00608

Frye, D., Zelazo, P. D., and Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cogn. Dev.* 10, 483–527. doi: 10.1016/0885-2014(95)90024-1

Gershoff, E. T. (2002). Corporal punishment by parents and associated child behaviors and experiences: a meta-analytic and theoretical review. *Psychol. Bull.* 128, 539–579. doi: 10.1037/0033-2909.128.4.539

Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action – performance of children 31/2-7 years old on a stroop-like day-night test. *Cognition* 53, 129–153. doi: 10.1016/0010-0277(94)90068-X

Henning, A., Spinath, F. M., and Aschersleben, G. (2011). The link between preschoolers' executive function and theory of mind and the role of epistemic states. *J. Exp. Child Psychol.* 108, 513–531. doi: 10.1016/j.jecp.2010.10.006

Horner, V., and Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan trogiodytes) and children (*Homo sapiens*). *Anim. Cogn.* 8, 164–181. doi: 10.1007/s10071-004-0239-6

Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Dev. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Hughes, C. H., and Ensor, R. A. (2009). How do families help or hinder the emergence of early executive function? *New Dir. Child Adolesc. Dev.* 2009, 35–50. doi: 10.1002/cd.234

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Dev. Cogn. Neurosci.* 1, 7–21. doi: 10.1016/j.dcn.2010.07.003

Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., et al. (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex* 46, 769–780. doi: 10.1016/j.cortex.2009.07.010

Kirkham, N. Z., Cruess, L., and Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Dev. Sci.* 6, 449–467. doi: 10.1111/1467-7687.00300

Kloo, D., and Perner, J. (2003). Training transfer between card sorting and false belief understanding: helping children apply conflicting descriptions. *Child Dev.* 74, 1823–1839. doi: 10.1046/j.1467-8624.2003.00640.x

Kloo, D., and Perner, J. (2005). Disentangling dimensions in the dimensional change card-sorting task. *Dev. Sci.* 8, 44–56. doi: 10.1111/j.1467-7687.2005.00392.x

Kochanska, G., Murray, K., and Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: from toddler to early school age. *Child Dev.* 68, 263–277. doi: 10.2307/1131849

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., and Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Dev.* 67, 490–507. doi: 10.2307/1131828

Landry, S. H., Miller-Loncar, C. L., Smith, K. E., and Swank, P. R. (2002). The role of early parenting in children's development of executive processes. *Dev. Neuropsychol.* 21, 15–41. doi: 10.1207/S15326942DN2101_2

Lehto, J. E., Juujarvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Lewis, C., and Carpendale, J. I. (2009). Introduction: links between social interaction and executive function. *New Dir. Child Adolesc. Dev.* 2009, 1–15. doi: 10.1002/cd.232

Lillard, A. S., and Else-Quest, N. (2006). The early years: evaluating montessori. *Science* 313, 1893–1894. doi: 10.1126/science.1132362

Lyons, D. E., Young, A. G., and Keil, F. C. (2007). The hidden structure of overimitation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19751–19756. doi: 10.1073/pnas.0704452104

Markus, H. R., and Kitayama, S. (1991). Culture and the self - implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224

McGuigan, N., Whiten, A., Flynn, E., and Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3-and 5-year-old children. *Cogn. Dev.* 22, 353–364. doi: 10.1016/j.cogdev.2007.01.001

Meins, E., Fernyhough, C., Wainwright, R., Clark-Carter, D., Das Gupta, M., Fradley, E., et al. (2003). Pathways to understanding mind: construct validity and predictive validity of maternal mind-mindedness. *Child Dev.* 74, 1194–1211. doi: 10.1111/1467-8624.00601

Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108

Moriguchi, Y. (2012). The effect of social observation on children's inhibitory control. *J. Exp. Child Psychol.* 113, 248–258. doi: 10.1016/j.jecp.2012.06.002

Moriguchi, Y., Evans, A. D., Hiraki, K., Itakura, S., and Lee, K. (2012). Cultural differences in the development of cognitive shifting: East–West comparison. *J. Exp. Child Psychol.* 111, 156–163. doi: 10.1016/j.jecp.2011.09.001

Moriguchi, Y., and Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6017–6021. doi: 10.1073/pnas.0809747106

Moriguchi, Y., and Hiraki, K. (2011). Longitudinal development of prefrontal function during early childhood. *Dev. Cogn. Neurosci.* 1, 153–162. doi: 10.1016/j.dcn.2010.12.004

Moriguchi, Y., and Hiraki, K. (2013). Prefrontal cortex and executive function in young children: a review of NIRS studies. *Front. Hum. Neurosci.* 7:867. doi: 10.3389/fnhum.2013.00867

Moriguchi, Y., and Hiraki, K. (2014). Behavioral and neural differences during two versions of cognitive shifting tasks in young children and adults. *Dev. Psychobiol.* 56, 761–769. doi: 10.1002/dev.21145

Moriguchi, Y., and Itakura, S. (2008). Young children's difficulty with inhibitory control in a social context. *Jpn. Psychol. Res.* 50, 87–92. doi: 10.1111/j.1468-5884.2008.00364.x

Moriguchi, Y., Kanda, T., Ishiguro, H., and Itakura, S. (2010). Children perseverate to a human's actions but not to a robot's actions. *Dev. Sci.* 13, 62–68. doi: 10.1111/j.1467-7687.2009.00860.x

Moriguchi, Y., Lee, K., and Itakura, S. (2007). Social transmission of disinhibition in young children. *Dev. Sci.* 10, 481–491. doi: 10.1111/j.1467-7687.2007.00601.x

Moriguchi, Y., Minato, T., Ishiguro, H., Shinohara, I., and Itakura, S. (2010). Cues that trigger social transmission of disinhibition in young children. *J. Exp. Child Psychol.* 107, 181–187. doi: 10.1016/j.jecp.2010.04.018

Moriguchi, Y., Okanda, M., and Itakura, S. (2008). Young children's yes bias: how does it relate to verbal ability, inhibitory control, and theory of mind? *First Lang.* 28, 431–442. doi: 10.1177/0142723708092413

Okanda, M., and Itakura, S. (2007). Do Japanese children say 'yes' to their mothers? A naturalistic study of response bias in parent-toddler conversations. [Empirical Study Quantitative Study]. *First Lang.* 27, 421–429. doi: 10.1177/0142723707081653

Perner, J., and Lang, B. (2000). "Theory of mind and executive function: is there a developmental relationship?" in *Understanding Other Minds: Perspectives from Autism and Developmental Cognitive Neuroscience*, eds S. Baron-Cohen, H. Tager-Flusberg, and D. Cohen (Oxford: Oxford University Press), 150–181.

Perner, J., Lang, B., and Kloo, D. (2002). Theory of mind and self-control: more than a common problem of inhibition. *Child Dev.* 73, 752–767. doi: 10.1111/1467-8624.00436

Qu, L. (2011). Two is better than one, but mine is better than ours: preschoolers' executive function during co-play. *J. Exp. Child Psychol.* 108, 549–566. doi: 10.1016/j.jecp.2010.08.010

Roskam, I., Stievenart, M., Meunier, J.-C., and Noël, M.-P. (2014). The development of children's inhibition: does parenting matter? *J. Exp. Child Psychol.* 122, 166–182. doi: 10.1016/j.jecp.2014.01.003

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., and Lee, K. (2006). The development of executive functioning and theory of mind - a comparison of Chinese and US preschoolers. *Psychol. Sci.* 17, 74–81. doi: 10.1111/j.1467-9280.2005.01667.x

Saxe, R., Schulz, L. E., and Jiang, Y. V. (2006). Reading minds versus following rules: dissociating theory of mind and executive control in the brain. *Soc. Neurosci.* 1, 284–298. doi: 10.1080/17470910601000446

Simpson, A., and Riggs, K. J. (2005). Inhibitory and working memory demands of the day-night task in children. *Br. J. Dev. Psychol.* 23, 471–486. doi: 10.1348/026151005x28712

Talwar, V., and Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Dev.* 79, 866–881. doi: 10.1111/j.1467-8624.2008.01164.x

Thorell, L. B., Lindqvist, S., Nutley, S. B., Bohlin, G., and Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12, 106–113. doi: 10.1111/j.1467-7687.2008.00745.x

Tomasello, M. (2009). *Why We Cooperate*, Vol. 206. Cambridge: MIT press.

Vygotsky, L. S. (1978). Mind In Society: *The Development of Higher Psychological Processes*. Cambridge, MA: Harvard Univ Press.

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684. doi: 10.1111/1467-8624.00304

Wiebe, S. A., Espy, K. A., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575

Zelazo, P. D., Frye, D., and Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cogn. Dev.* 11, 37–63. doi: 10.1016/S0885-2014(96)90027-1

Zelazo, P. D., Müller, U., Frye, D., and Marcovitch, S. (2003). The development of executive function in early childhood. *Monogr. Soc. Res. Child Dev.* 68, Serial No. 274.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2014; accepted: 14 April 2014; published online: 29 April 2014.
Citation: Moriguchi Y (2014) The early development of executive function and its relation to social interaction: a brief review. Front. Psychol. ***5****:388. doi: 10.3389/fpsyg.2014.00388*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Moriguchi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 27 May 2014
doi: 10.3389/fpsyg.2014.00487

Relations between executive function and emotionality in preschoolers: Exploring a transitive cognition–emotion linkage

David E. Ferrier*, Hideko H. Bassett and Susanne A. Denham

Department of Psychology, George Mason University, Fairfax, VA, USA

Edited by:
Philip D. Zelazo, University of Minnesota, USA

Reviewed by:
Gary Morgan, City University London, UK
Ruth Ford, Anglia Ruskin University, UK

****Correspondence:***
David E. Ferrier, Department of Psychology, George Mason University, 4400 University Drive, Fairfax, VA 22030, USA
e-mail: dferrier@masonlive.gmu.edu

Emotions play a crucial role in appraisal of experiences and environments and in guiding thoughts and actions. Moreover, executive function (EF) and emotion regulation (ER) have received much attention, not only for positive associations with children's social–emotional functioning, but also for potential central roles in cognitive functioning. In one conceptualization of ER (Campos et al., 2004), processes of ER, and those of emotional expression and experience (hereafter referred to as *emotionality*) are highly related and reciprocal; yet, there has been little research on young children's EF that focuses on emotionality, although it is easily observed within a classroom. The two goals of the study were to: (1) investigate the relatively unexplored role of emotionality in the development of EF in early childhood and (2) assess the relations between an observational rating of EF obtained after direct assessment with a standardized EF rating scale. We predicted that observed emotionality and EF would both demonstrate stability and predict one another within and across time. 175 children aged 35–60 months were recruited from Head Start and private childcare centers. Using partial least squares modeling, we chose T1 emotionality as the exogenous variable and tested pathways between emotionality and EF across two time points, 6 months apart. Results showed that both T1 observed EF and emotionality predicted their respective T2 counterparts, supporting the idea that both constructs build upon existing systems. Further, T1 emotionality predicted T1 observed EF and the T2 BRIEF-P composite. In turn, T1 observed EF predicted emotionality and the T2 BRIEF-P composite. These findings fit with literature on older populations in which EF and emotionality have been related, yet are the first to report such relations in early childhood. Last, T1 observed EF's positive prediction of the T2 BRIEF-P composite lends credence to the use of both EF measures in applied and research settings.

Keywords: executive function, preschool, emotional expression, emotion regulation, self-regulation

INTRODUCTION

Emotions are thought to play a crucial role in our appraisal of experiences and environments, in guiding our thoughts and actions, as well as regulating our behavior, and in adapting to situations (Cole et al., 2004; Lehtonen et al., 2012). Whereas researchers have started recognizing the interconnections between emotion and cognition, particularly between executive functions (EFs) and emotion regulation (ER; e.g., Blair, 2002; Blair and Diamond, 2008; Blankson et al., 2013), little research with young children has been focused on other aspects of emotion such as emotional expression, even though it is easily observed within a classroom context. In this study, we examine the role of emotional expression and experience (hereafter referred to as *emotionality*) and its interconnection with the development of executive functioning. Before moving to our main questions, however, we should examine the literature already existing on EF and ER.

Executive function and ER abilities have received a large amount of attention for not only their associations with benefits in children's social–emotional functioning, but also for their suggested critical roles in cognitive functioning (Denham, 2006; Bassett et al., 2012). Moreover, both EF and ER are considered to be aspects of self-regulation (Smith-Donald et al., 2007; Jahromi and Stifter, 2008), which we believe encompasses an individual's ability to control one's emotional, behavioral, and cognitive actions and responses (Smith-Donald et al., 2007; Jahromi and Stifter, 2008).

To further define these two aspects of self-regulation, EF is considered a collection of higher-order brain functions, generally viewed as incorporating working memory, attention shifting, and inhibitory control (Miyake et al., 2000; Garon et al., 2008). In terms of its importance, Riggs et al. (2006) wrote of the connections between EF and numerous correlates of social–emotional functioning, such as theory of mind and delay of gratification. Additionally, positive academic achievement outcomes have also been linked to greater EF abilities (e.g., Blair and Razza, 2007; Bierman et al., 2008).

Although different working definitions exist for ER, Campos et al. (2004) chose to define ER as any alteration in the system responsible for the generation and behavioral manifestation of emotions. More specifically, it has been considered "the process of

initiating, maintaining, modulating, or changing the occurrence, intensity, or duration of internal feeling states and emotion-related physiological processes, often in the service of accomplishing one's goals" (Eisenberg et al., 2000, p. 137; see also Thompson et al., 2008). Research has shown that children who have trouble regulating their emotions in the classroom are more prone to exhibit later psychopathology (e.g., Cole et al., 1996), and aggression (e.g., Calkins and Marcovitch, 2010), as well as to suffer from peer rejection, increased anhedonia about school, and poor academic outcomes (Trentacosta and Izard, 2007; Ursache et al., 2012). Further, there is empirical support for the role that ER plays in promoting more positive attributes, such as social competence (Denham et al., 2003) and school adjustment (Herndon et al., 2013).

Clearly both abilities have important sequelae. But how do we view their interrelation? Consistent with the view that ER and cognitive regulation (i.e., EF) are both narrow domains of the broader self-regulation construct (Smith-Donald et al., 2007; Jahromi and Stifter, 2008), Ursache et al. (2012) propose that the connections between the self-regulatory aspects of ER and EF are reciprocal in nature.

Consider the literature on infants which, within the past decade, have both suggested that cognition and emotion are dynamically interwoven (Bell and Wolfe, 2004) and that early indicators of ER positively predicted later EF ability at age four, in children high in emotional reactivity (Ursache et al., 2013). Additional research has provided support for behavioral assessments and parental ratings of inhibitory control in young children concurrently predicting their ER abilities (Carlson and Wang, 2007). Other research investigating parental ratings of ER, suggested that ER supports the later development of EF in preschool-aged children (Blankson et al., 2013). Viewed through a wider lens, findings from studies such as Blankson et al. (2013) and Carlson and Wang (2007) support a transactional model between both EF and ER (Ursache et al., 2012).

These relations are also consonant with developmental neuroscience research, which has also suggested a deeper connection between cognition and emotion centers of the brain (e.g., Cacioppo and Berntson, 1999; Bell and Wolfe, 2004; Carlson and Wang, 2007). Although developmental cognitive neuroscience studies offer suggestions of cognition–emotion linkages, a prevailing notion about the relation between ER and EF suggests that the corresponding areas of the brain connected to these functions are neurologically similar. Calkins and Marcovitch (2010) wrote that empirical connections between EF and ER are, in part, due to areas that are active in the prefrontal cortex (PFC) of the brain. Specifically, two subdivisions within the anterior cingulate cortex (ACC) of the PFC are responsible for cognitive and attentional processes and emotional processes, respectively. In agreement with views from Davidson et al. (2000), Denham et al. (2012a) and Ursache et al. (2012), the model proposed by Calkins and Marcovitch (2010) also suggests that the relations between these two subdivisions of the ACC are reciprocal in nature.

Whatever processes account for this reciprocity, its existence implies that the development, whether typical or atypical, of one aspect of a child's regulatory capabilities affects the trajectory of other self-regulatory processes. Thus, testing the relations between EF and other aspects of emotion should aid developmental science in understanding equally relevant regulatory processes. In turn, integrating across specific research niches (i.e., EF, ER, and emotionality; Duncan, 2012) can be useful in constructing a more unified knowledge base aimed at preventing specific self-regulatory deficits from cascading across social, emotional, cognitive, and academic domains (see also Blair et al., 2005).

Thus, whereas the interplay between EF and ER is empirically supported within early childhood, the contribution of emotional expression has been overlooked in the self-regulatory literature (Blair, 2002; Riggs et al., 2006; Blair and Razza, 2007; Bierman et al., 2008; Brock et al., 2009). Studies examining cognition–emotion connections have mainly focused on the relation between cognitive and ER (e.g., Calkins and Marcovitch, 2010; Iida et al., 2011); however, a new conceptualization of ER may be what is needed to rectify this limitation of earlier research. In this new formulation, the processes of ER and those of emotionality are highly related, often co-occurring, and reciprocal (Campos et al., 2004). This conceptualization is central to our attempts to address the unanswered relations between EF with emotionality.

More specifically, although the two-factor approach of ER, in which the processes of emotionality and ER are distinguished, has been widely accepted in the past, this model may be an oversimplification (Cole et al., 2004). Instead, uniting emotionality and ER in a one-factor model is a fruitful alternative because it may more faithfully depict the actual process of emotion (Campos et al., 2004). That is, emotions are expressed and experienced almost simultaneously with their regulation; in fact, much of the difficulty in defining and measuring ER lies in its inseparability from emotionality.

Considering the key role that such emotionality plays in ER, then, one would anticipate emotionality, examined uniquely, to also both affect and be affected by the developmental progression of other self-regulatory processes, namely, by an individual's EF abilities, just as are an individual's ER abilities. Thus, the overarching goal of the present study is to examine this yet relatively unexplored connection between cognition and emotion: the relation between preschoolers' EF and emotionality. Finding the relation between EF and emotionality will have a significant benefit not just in research community but also in applied settings. Because, unlike direct assessments of ER that usually involve standard lab procedures eliciting negative emotions from children to observe how they regulate the emotions, emotionality is easily observed in natural settings (e.g., classroom) by preschool teachers.

Based on Campos et al. (2004) unitary process of ER and emotion, we hypothesized that emotionality would be related to the development of EF, and that over time, a reciprocal function between EF and emotionality would be found. Falling in line with the developmental neuroscience literature, we draw additional support for our position from the idea that the more mature portions of the brain responsible for negative emotionality (e.g., amygdala) are capable of inhibiting the deployment, and development, of executive cognitive processes housed in later maturing areas (e.g., PFC; Blair, 2002).

Although research examining the relations between emotionality and EF is scarce with young children, empirical support has been provided for the emotionality-EF link from research with adolescents/young adults. For example, poor EF was found to be related to an increased tendency to express negative affect in college students (Bridgett et al., 2013). In functional neuroimaging research with college students, Luu et al. (2000) also found that affective distress was closely related to frontal lobe EF. If emotionality and higher-order cognitive regulation (i.e., EF) are related in adults, then, examining the relations of these constructs in young children will further aid our understanding of the emotion–cognition interconnectivity from a developmental perspective.

A secondary goal of this paper is to examine the relations between an observational rating of EF obtained after direct assessment with a standardized rating scale. This goal is in order because of difficulties with specificity of EF assessments across age (Best and Miller, 2010). Considerable research has exemplified the range of growth that occurs during the preschool years in young children's EF (Hughes, 1998; Garon et al., 2008). A common theme amongst prior research was the prediction that measuring EF in preschool-aged children would be difficult due to rapid development, yielding tasks either too easy, resulting in ceiling effects, or tasks too difficult, yielding significantly negatively skewed findings (Hughes, 1998; Isquith et al., 2004; Blair et al., 2005; Carlson, 2005; Garon et al., 2008; Bassett et al., 2012). With the growing notion that inhibitory control and sustained attention not only act as rudimentary forms of EF (Carlson and Wang, 2007; Jahromi and Stifter, 2008; Graziano et al., 2011; Blankson et al., 2013), but also are implicated in the development and utilization of ER, careful measurement and examination of these constructs in a preschool population is of key importance (Riggs et al., 2006).

Two studies have recently contributed to solving this issue of age effects in measuring preschool-aged children's EF, by using ratings rather than direct assessment. Smith-Donald et al. (2007) developed a two-part assessment of self-regulation, the Preschool Self-Regulation Assessment (PSRA), which is composed of a direct assessment battery and an assessor report (AR) capturing global behavior. The AR consists of several rating items from the Leiter-R social–emotional rating scale (Roid and Miller, 1997) and the Disruptive Behavior-Diagnostic Observation Schedule (Wakschlag et al., 2005). A second study conducted by Isquith et al. (2004) sought to downwardly shift the Behavior Rating Inventory of Executive Function (BRIEF) for a preschool sample (BRIEF-P).

Together, the AR and BRIEF-P have provided measures that do not fluctuate with age as do the more commonly used performance-based tasks, and allow for a more generalizable view of EF. To date, however, there have been no studies looking at relations between the AR and the BRIEF-P. Investigations into their relation could bolster the utilization of rating scales, particularly of scales that are of relative ease of use and do not require a great expense of time.

In sum, research has demonstrated a connection between ER and EF (e.g., Carlson and Wang, 2007). Especially in young children, however, EF's relation to other aspects of emotion has not been explored. This new unitary perspective on emotionality and ER impels us to consider the heretofore little explored linkage of preschoolers' emotionality and EF.

In the present study, we collected data using multiple methods and reporters at two time points, to enable us to study relations across short-term longitudinal periods. Specifically, trained research assistants observed children's emotional expression in naturalistic settings, rated their cognitive regulation (i.e., EF) based on observations of their behaviors during several direct assessments (i.e., social and emotional competence and school readiness), and preschool teachers completed a standardized questionnaire assessing preschoolers' EF.

Thus, as our first problem question, we examined the relations between emotionality and EF both within and across time in a multi-method approach; we would expect each to show continuity across time, and for emotionality to contribute positively to EF. Although we believe that there is a transactional reciprocity between EF and emotionality, consistent with others (Ursache et al., 2012), with a preschool age sample, we chose emotionality to initially serve as an exogenous variable given that areas of the brain responsible for emotion tend to reach maturation earlier than areas responsible for EF (Martel, 2009; Nigg et al., 2010; Kanske and Kotz, 2012). For this reason, we are testing the directional pathway from emotionality to EF in early childhood within each time period, with cross-lagging pathways between both EF and emotionality between time periods (see **Figure 1**).

Second, given our focus on early childhood development and education, we wished to see how teachers' views of end-of-year EF were predicted by earlier and concurrent observed emotionality and EF; triangulating across these indices strengthens claims for validity, and thus usefulness, of the teacher ratings of EF in research and applied settings.

METHOD

PARTICIPANTS

The current sample is part of a multi-year, multi-site larger project investigating the impact and role that preschool teachers play in facilitating social–emotional competencies. Participants were recruited from ten local Head Start programs and private child-care facilities in the surrounding northern Virginia area, and were culturally, socio-economically, and racially diverse. Children participating were identified via parent contact at recruitment events held at child pick-up, information sessions held at the facilities, and/or through the help of facility social workers and directors.

One hundred seventy-five children aged 35–60 months were recruited for this study and parental consent was attained. Of these, complete data was obtained for 143 (81%) children. Additionally, 36% ($N = 52$) of children were from federally funded Head Start programs. Females comprised slightly more than half of the sample (52.4%). Parents who provided demographic information self-identified as 43.4% Caucasian, 13.9% African–American, 4.9% Asian, 4.2% Multiracial, 3.5% Other; 30.1% of parents did not report their child's race. Hispanics/Latinos constituted 11.2% of the sample; 28.7% of parents did not report their child's ethnicity.

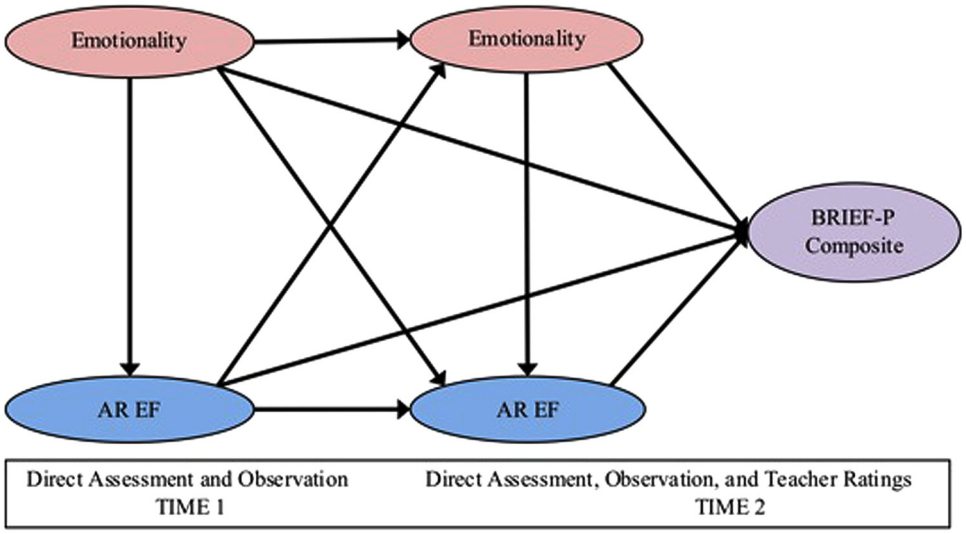


FIGURE 1 | Partial least-squares outer model.

PROCEDURES

Assessments comprised of observation systems and rating scales. Children were assessed in the fall and ∼6 months later in the spring. Trained research assistants were either graduate or undergraduate students or volunteers who had extensive training to ensure reliability and appropriate assessment techniques. Because this study is part of a larger grant, additional measures, unrelated to the current study, were administered to participants investigating their social–emotional development. Three direct assessments were administered in a quiet testing environment at the schooling facility at both time points; these measured school readiness, emotion knowledge, and social problem-solving. Following each of these three sessions at both time points, research assistants completed a rating scale about observed EF behavior specific to that session. Additionally, children's emotionality was observed four times in both the fall and spring data collection periods. Teachers completed a rating scale in the spring session assessing EF in real-world contexts.

To thank the child for participating, a small gift (e.g., small box of crayons or small vial of bubbles) was given to the child at the end of each assessment period. Teachers were compensated $15 for the completion of rating scales for each child.

MEASURES OF PRESCHOOLERS' EMOTIONALITY AND EXECUTIVE FUNCTIONING

Minnesota Preschool Affect Checklist – Revised/Shortened (MPAC-R/S; Denham et al. 2012b)

The MPAC-R/S is an 18-item observational measure of social–emotional behavior. Previous research has shown that the MPAC-R/S observation system is a valid and reliable tool, with emotionality and regulation related to later preschool classroom adjustment, as well as classroom adjustment and academic success in kindergarten, even age, gender, and prior school success controlled (Denham, 2006; Herndon et al., 2013).

Four 5-min observations were completed by trained observers in both the fall and spring of the academic year and were collected during periods of recess, freeplay, and activity station ("centers") times. Attempts were made to vary the contexts in which the MPAC-R/S captured data to reduce situation-specific factors from reducing validity. Furthermore, MPAC-R/S sessions were collected on separate days to allow for variability.

In this study, five items were used to specifically focus upon and assess children's positive and negative emotional expression [e.g., "The child displays positive affect in any manner (i.e., facial vocal, or bodily affect)," and "The child directs negative affect specifically at a particular person when already in contact with them"]; coders take note only of directly observable emotional expressiveness, and, although it is impossible to determine whether any individual child was exerting any internal regulation during any one individual observation period, we feel that by collapsing over several occasions these items are good indicators of emotionality. In analyses to follow, differences in standard scores for positive and negative expression indicated emotionality.

Further, the MPAC-R/S allows for observation of behavioral evidence of ER and dysregulation. Thus, in this study, indices for positive regulation (focusing solely on using language to regulate negative emotion) and dysregulation (focusing on venting outbursts) were also included for subsidiary analyses.

Minnesota Preschool Affect Checklist – Revised/Shortened item content, as well as internal consistency information for the indices of emotionality and regulation/dysregulation, can be seen in **Table 1**. Inter-observer reliability for these data was obtained by calculating averaged measure intraclass correlations (ICCs) for the group of observers, including a master coder. Across two training periods, ICCs were 0.94 and 0.95 for positive emotional expression, 0.97 and 0.98 for negative emotional expression, 0.87 and 0.74 for positive regulation, and 0.98 and 0.99 for dysregulation.

Table 1 | MPAC-R/S observation items.

Positive emotion ($\alpha = 0.77$ and 0.67 for T1 and T2, respectively)
1. The child displays positive emotion in any manner (i.e., facial, vocal, or bodily emotion). The child's behaviors must match the context of a given situation. Examples: smiling, laughing, singing, dancing, etc.
2. The child directs positive emotion specifically at a particular person when already in contact with them. Emotion is directed at a specific person.
3. The child displays positive emotion when in a social situation but does not direct it to anyone in particular.
Negative emotion ($\alpha = 0.92$ and 0.93 for T1 and T2, respectively)
1. The child displays negative emotion in any manner (i.e., facial, vocal, or bodily emotion). The child's behaviors must match the context of a given situation.
2. The child directs negative emotion specifically at a particular person when already in contact with them. Emotion is directed at a specific person.
Emotion regulation: positive reactions to emotionally arousing problem situations ($\alpha = 0.79$ and 0.80 for T1 and T2, respectively)
1. The child promptly verbally expresses the feelings arising from a problem situation, then moves on to the same or a new activity (versus withdrawing, displacing the emotion onto others or objects, or staying upset).
2. The child shows primarily neutral or positive emotion during this behavior.
Emotion dysregulation: negative reactions to emotionally arousing problem situations (usually anger-related; $\alpha = 0.37$ and 0.59 for T1 and T2, respectively)
1. The child displays context-related interpersonal aggression (verbal or physical). Someone does something emotionally arousing, to which the child responds with aggression (emotionally arousing preceding event must be observed).
2. The child hits, kicks, shoves, knocks over, or throws objects (emotionally arousing preceding event must be observed).

Average inter-item correlations for all scales were significant (Spiliotopoulou, 2009).

Assessor report

The AR, adapted from a measure originally compiled by Smith-Donald et al. (2007), consists of 12 items asking the researcher to assess the child's emotional expression, attention, and behavior over the course of an assessment interaction in which data was collected. All items are rated on a 4-point Likert scale ranging from 0 to 3, with five items reverse-coded to reduce acquiescence bias. The AR was administered following direct assessments not in this study at three time points and scores were aggregated to consolidate data into two variables, fall (T1) and spring (T2). Although the AR consists of six scales (Confidence, Affective Balance, Engagement, Attention, Emotion regulation, and Inhibition), only the Attention and Inhibition scales were used in the current study. An example of a prompt assessing Attention was "Distracted by sights and sounds throughout assessment period," and an Inhibition prompt was "Lets examiner finish before starting task; does not interrupt," examiners then rate the frequency and intensity from 0 to 3.

In terms of reliability, internal consistency values for the AR factors of Attention (six items) were $\alpha = 0.77$ at T1 and $\alpha = 0.74$ at T2, and for Inhibition (three items), were $\alpha = 0.54$ for T1 and $\alpha = 0.61$ for T2. Because having a small number of items can negatively impact alpha values, examining the mean inter-item correlations can also provide an accurate representation of internal consistency (Clark and Watson, 1995; Spiliotopoulou, 2009). Mean inter-item correlations for AR Attention were 0.35 at T1 and 0.33 at T2, $ps < 0.001$. For Inhibition, corresponding correlations were 0.29 for T1 and 0.34 for T2, $ps < 0.001$. These values suggest that these items are appropriately related. For inter-observer reliabilities, averaged measure ICC was 0.98 for both Attention and Inhibition.

In terms of validity for the scales utilized here, analyses of the AR by the original authors (Smith-Donald et al., 2007) reported that there were non-significant gender differences, suggesting the presence of construct validity. Furthermore, Smith-Donald et al. (2007) provided concurrent validity for the original AR, showing significant correlations between their Attention/Impulse Control factor and both externalizing and internalizing problems, as well as social competence.

Behavior Rating Inventory of Executive Function – Preschool Version (BRIEF-P; Gioia et al. 2003)

Teachers were asked to complete the BRIEF-P at the end of the data collection cycle in the spring of the academic year. The BRIEF-P is a standardized rating scale providing information about the executive functioning of children from ages 2 to 5 years. The measure consists of 63 items providing five distinct scales, one composite scale and three overlapping summary indexes. The BRIEF-P yields five scales assessing Inhibitory Control, Attention Shifting, Emotional Control, Working Memory, and Plan/Organize. These scales reflect all facets of the larger construct of EF and permit comparative benchmarks in EF between subjects. In total, the BRIEF-P takes approximately 10 min to complete.

Excellent internal consistency was found for the five scales (Shift, $\alpha = 0.90$; Inhibition, $\alpha = 0.95$; Working Memory, $\alpha = 0.95$; Emotional Control, $\alpha = 0.93$; Plan/Organize, $\alpha = 0.90$). These values were highly similar to the reported values from the test authors (Gioia et al., 2003). Reported validity for the BRIEF-P

demonstrated significant correlations across many scales on the Behavior Assessment System for Children – Parent Rating Scales (BASC) with correlations ranging from −0.83 to 0.76 in expected directions.

DATA ANALYSIS

Partial least squares modeling (PLS: Falk and Miller, 1982; Ringle et al., 2005) was utilized to answer our major problem questions. In common with other modeling techniques, a measurement (outer) model as well as a structural (inner) model is specified. For the outer model, PLS estimates latent variables (LVs) based on the shared variance of the manifest variables, using principal components weights of the manifest variables. As such, each indicator varies in how much it contributes to the LV, resulting in the best possible combination of weights for predicting the LV while accounting for all manifest variables, a distinct advantage of the method (Tsethlikai, 2010).

This method, which is becoming more widely known by developmentalists (e.g., Brody et al., 1994; Cowan et al., 1996; Marjoribanks, 1997; Davies and Cummings, 1998; Isley et al., 1999; Denham et al., 2002, 2003; Bronstein et al., 2005; Tsethlikai, 2010, 2011), also allows exploration of hypothesized relations among constructs without some of the restrictions of LISREL structural modeling techniques. In particular, PLS is appropriate for use with relatively small groups of participants, although it does require a reasonable LV: participant ratio (e.g., 10 times the number of manifest variables for the LV with the largest number of manifest variables, or 10 times the largest number of paths directed at a LV; Henseler et al., 2009). Further advantages include its lack of stringent assumptions such as those regarding observational independence and normality of residuals (Marjoribanks, 1997), as well as error-free measurement (Tsethlikai, 2011).

Outer measurement models provide information on the psychometric reliability of our constructs' LVs. Inner measurement models do not allow for bidirectional pathways (Barroso et al., 2010), thus, only a unidirectional pathway between LVs was tested within each time point. This estimation assessed predictive validity via the relations among LVs and significant, hypothesized paths. Bootstrapping procedures then allow for significance testing of each path. Further, both inner and outer measurement models provide information on discriminant validity, when LV correlations are compared to the square root of the LV's average variance extracted (AVE). For this study, LVs are as follows: for both T1 and T2: emotionality and AR EF, and for T2 only: the BRIEF-P Composite. In our model, manifest variables (indicators) were hypothesized to form these LVs, and all hypothesized paths among these LVs were of interest (see **Figure 1**).

RESULTS

OUTER MODEL

Using Smart-PLS^{TM} (Ringle et al., 2005), we first examined acceptability of the outer measurement model. Regarding the outer model, three criteria are present: (a) the set of manifest variables represents the same underlying construct (AVE), with a reasonable total explained variance (R^2); (b) the manifest variables also form an internally consistent LV (composite reliability); and (c) each manifest variable loads sufficiently on its LV to support its retention (i.e., each manifest variable contributes to its LV and represents the construct in a similar manner as other manifest variables). According to Henseler et al. (2009), composite reliabilities for all LVs formed by the hypothesized collection of manifest variables should be $\geq$0.60, and AVE should be $\geq$0.50. Finally, each manifest variable's outer model loading should be $\geq$0.70.

Findings for our model suggested the following (see **Table 2**): (a) all composite reliabilities were >0.60 and (b) all AVEs were >0.50. Further, all manifest loadings were >0.70. Thus, the outer model met all criteria so that inner model analyses could proceed.

CONVERGENT AND DISCRIMINANT VALIDITY

Table 3 shows the square roots of the AVEs and the correlations amongst LVs. This information can yield information on both convergent and discriminant validity. First, for convergent validity, a LV should explain better the variance of its own indicator than that of other LVs. One way to determine this point is to compare the square root of each LV's AVE with all correlations involving that LV. If the correlation between any two LVs is less than the square root of either of their individual AVE's, this suggests that each has more internal (extracted) variance than variance shared between the LVs.

Second, if these criteria are met for a target LV and *all* the other LVs, this suggests the discriminant validity of the target LV (Fornell and Larcker, 1981). Correlations with other LVs of less than |0.7| are also frequently accepted as evidence of discriminant validity. The information in **Table 3** shows that these criteria for both convergent and discriminant validity are met for all LVs in the model. Finally, examination of cross-loadings indicated that each manifest variable's loading was far higher for its assigned LV than the other LVs; by this criterion as well (not tabled), these LVs showed good discriminant validity.

INITIAL EVALUATION OF THE INNER MODEL

Given these validity results, we can continue to an examination of the inner model. The first step here is to examine the LVs' correlations in respect to hypothesized relations among them. As can be seen in **Table 3**, MPAC-R/S Emotionality showed T1 to T2 stability, and both time points' index of emotionality was related to the BRIEF-P Composite. T2 Emotionality was also related to observed EF at both time points. Finally, AR EF showed T1 to T2 stability, and each time point's index of observed EF was related to the BRIEF-P Composite.

OVERVIEW OF STRUCTURAL PATH MODEL

Figure 2 depicts the final structural model. Path coefficients in the model can be interpreted as standardized beta weights, each estimated after all other paths' effects have been controlled. To assess whether the paths were significant, bootstrapping resampling (Efron and Gong, 1983) was performed. In this procedure, the PLS parameters of a series of random subsamples of the total sample are iteratively tested, until significance can be estimated based on their convergent findings.

Our final structural model can be summarized by noting the following significant direct effects of LVs: (1) T1 Emotionality

Table 2 | Outer model and final R^2s for latent variables.

LV	LV			Manifest loading
Manifest variable	**AVE**	R^2	**Composite reliability**	
Time 1		—	—	
Emotionality	1.00		1.00	
MPAC-R/S emotionality				1.00
Executive function: AR EF	0.72	0.026	0.84	
AR attention				0.868
AR inhibitory control				0.835
Time 2				
Emotionality	1.00	0.149	1.00	
MPAC-R/S emotionality				1.00
Executive function: AR EF	0.803	0.328	0.89	
AR attention				0.917
AR inhibitory control				0.875
Executive function:	0.705	0.150	0.92	
BRIEF-P composite				
Emotional control				0.819
Inhibitory control				0.876
Planning and organization				0.884
Shifting				0.736
Working memory				0.873

AVE = Average variance extracted.

Table 3 | Inner model latent variable correlations.

Scale and time point	1.	2.	3.	4.	5.
1. Emotionality T1	**1.00**				
2. Emotionality T2	0.28*	**1.00**			
3. AR EF T1	0.16^{+}	0.31**	**0.85**		
4. AR EF T2	0.05	0.17*	0.57***	**0.89**	
5. BRIEF-P composite T2	0.24*	0.23*	0.32**	0.22*	**0.84**

*Square roots of AVEs appear in bold on the diagonal; LV correlations appear below the diagonal. $^{+}p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.*

predicted T1 AR EF, T2 Emotionality, and BRIEF-P Composite. T1 AR EF predicted T2 Emotionality, as well as T2 AR EF, and BRIEF-P Composite. T2 AR EF also predicted the BRIEF-P Composite.

SUBSIDIARY ANALYSES

Two further PLS analyses were undertaken. In the first, several iterations of PLS were attempted. Regulation and dysregulation were included along with emotionality, to show that emotionality was in fact key in the model, rather than merely a marker of regulation. However, outer loadings for regulation and dysregulation in this model did not meet the standard of 0.70 for continued inclusion in the LV. Strict PLS modeling would then require reverting back to the model in **Figure 2**. In these analyses, however, the outer loading for dysregulation, was >0.60 at both T1 and T2, so that a model with emotionality and dysregulation was performed. It was virtually identical for that including only emotionality, suggesting that in fact observed emotionality is key in these analyses. Hence, the primary findings for our research question regarding emotionality and EF, as noted in **Figure 2**, remain.

Second, our research question on how teachers' views of end-of-year EF are predicted by earlier and concurrent observed emotionality and EF was refined methodologically by deleting the Emotional Control scale from the BRIEF-P LV, to make an even purer EF construct. Again, the PLS model was almost identical to that in **Figure 2**, suggesting that the original BRIEF-P LV, which is based on psychometric standardization of the measure, can be retained for discussion.

DISCUSSION

OVERVIEW

This research describes an original endeavor to investigate the relations between emotionality observed in natural settings (i.e., while interacting with peers in preschool classroom) and EF in a preschool population. Conceptualizing ER and emotionality to use the same processes, based on the framework proposed by Campos et al. (2004), we expanded our focus to specifically examine whether relations between EF and emotionality were

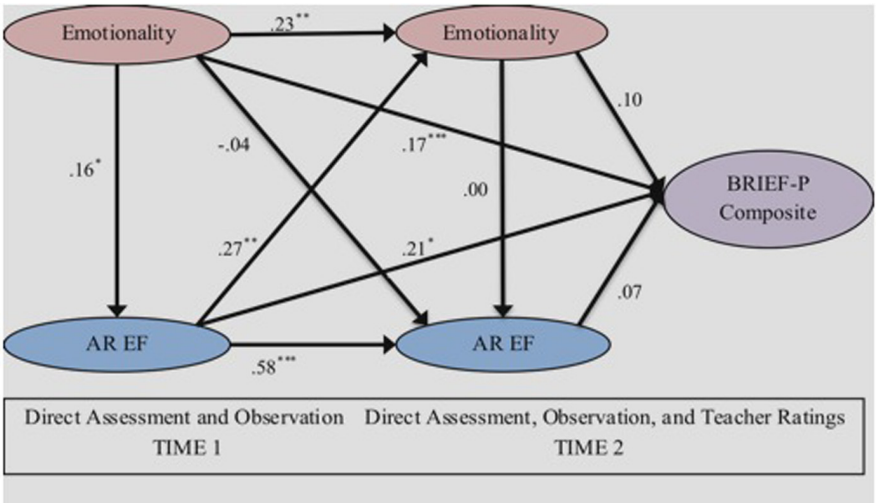


FIGURE 2 | Partial least-squares inner model. *Note.* Path coefficients may be interpreted as standardized beta coefficients. Levels of significance determined by *t*-values from bootstrapping procedures and may vary according to the standard error of the path coefficient; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

present as have been found repeatedly between EF and ER. Over time, we believe that emotionality and EF will become reciprocal, a position supported by others (e.g., Blair, 2002). However, given the statistical procedure used and paired with research that has posited that emotion processes develop earlier (Nelson, 1994; Blair, 2002), and in turn influence, more complex cognitive processes, (i.e., EF; Calkins and Marcovitch, 2010; Nigg et al., 2010; Blankson et al., 2013; Ursache et al., 2013), we predicted that measures of emotionality would in turn predict later EF. Using PLS modeling, we were able to test our proposed pathway between emotionality and EF across two time points, approximately 6 months apart; emotionality at T1 predicted observed AR EF at that time, as well as the T2 BRIEF-P Composite. AR EF at T1, in turn, predicted emotionality at T2, as well as the T2 BRIEF-P Composite.

ANOTHER EMOTION–COGNITION LINKAGE: EF AND EMOTIONALITY

Our primary goal to examine the continuity of EF and emotionality across two time points and to examine the contribution of emotionality to later EF development was supported by our current findings. Subsidiary analyses, (1) including observed dysregulation and (2) excluding the Emotional Control scale from the BRIEF-P LV, did not yield different results from our proposed model. Thus, we are confident to conclude that a significant relation exists between preschoolers' emotionality and EF. Implications from these findings contribute to the growing literature stressing the importance of emotions in preschoolers' optimal development (e.g., Denham, 2006; see also Chaplin and Aldao, 2013). Although these findings do not neurologically examine whether portions of the brain dealing with emotion development underlie those areas responsible for EF, the results lend support to previous models detailing their interconnection (Calkins and Marcovitch, 2010; Ursache et al., 2013). Further, this research serves to emphasize that emotionality is implicated in EF abilities, just as ER is often found relating to EF (e.g., Blankson et al., 2013), which suggests that emotionality and ER are part of a larger interconnected self-regulatory network. Finding that T1 scores of EF and emotionality predicted their T2 counterparts also supports the idea that both EF and emotionality are constructs that build upon existing systems (Denham, 2007; Garon et al., 2008). These findings fit with existing literature looking at older populations in which EF and emotionality have also been related (Luu et al., 2000; Bridgett et al., 2013), yet are the first to examine such relations in early childhood.

Thus, the current study contributes empirical support for the promotion of both positive emotionality and EF in preschoolers. In recent times, there have been numerous studies that have separately showcased advantageous outcomes associated with positive emotionality and early precursors of self-regulatory processes, including EF and ER (Denham et al., 2003, 2013; Denham, 2006; Riggs et al., 2006; McClelland et al., 2007; Liew, 2012). Having adequate EF and ER skills and manifesting a more positive emotionality is often considered critical for ensuring numerous positive outcomes, such as school readiness and social–emotional competence (e.g., Denham et al., 2003, 2012a, 2013; Denham, 2006; Trentacosta and Izard, 2007; Brock et al., 2009; Ursache et al., 2012; Herndon et al., 2013). Demonstrating that emotionality contributes to later EF should, we hope, serve to increase the importance of both emotions and EF abilities within the preschool classroom.

Conversely, deficits in EF, ER or more negative emotionality may lead to negative outcomes that could adversely affect numerous facets of optimal development across domains (Denham et al., 2003, 2012a; Denham, 2006; Bassett et al., 2012). This assertion was supported by the current findings, as greater negative emotionality (i.e., indexed by lower or negative emotionality scores) predicted greater EF problems on the BRIEF-P. Through the lens of an educational administrator, these children with greater negative emotionality and/or lower EF would

require additional time, effort, and resources from teachers, parents, and supportive staff if problematic behavior were being exhibited.

Developmental researchers are increasingly engaged in addressing and understanding precursors of developmental problems, particularly attention-deficit hyperactivity disorder (ADHD; e.g., Barkley, 1997; Anastopoulos et al., 2011). Children diagnosed with ADHD are marked by lower levels of EF, which have been linked with problems in emotional competence, specifically, ER (e.g., Barkley, 1997; Blair et al., 2005). Understanding early contributing factors to EF will aid preventative literature.

Further, research has shown that exhibiting greater negative emotionality has been strongly linked to numerous poor outcomes, particularly in the preschool and early formal schooling years (Belsky et al., 2001; Denham, 2006; Anastopoulos et al., 2011). Previous research has shown that outcomes such as high ratings of negative behavior by the classroom teacher (Herndon et al., 2013) and lower sociometric likeability and teacher ratings of social competence (Denham et al., 2003) are related to negative emotionality and emotion dysregulation. Recently, a push for preventative practice has underscored the importance of addressing such emotional competence deficits (see also Izard, 2002).

RELATION BETWEEN THE ASSESSOR REPORT AND THE BRIEF-P

As many teachers are becoming overburdened by high-stakes testing requirements, the utility of easy-to-use assessment measures trumps those that are more laborious and time-consuming. Thus, a second aim of this study was to provide evidence of the BRIEF-P's usability in research and applied settings. Although rating scales of EF typically manifest low to moderate correlations with direct assessments of the same constructs they are both said to measure, rating scales are less context-specific, averaging the rater's evaluation of the child over many observations. This property of rating scales has led to the view that they may accurately capture real-world portrayals of EF development (Cairns and Green, 1979; Isquith et al., 2004). Furthermore, the ease of rating scales eliminates the need for extensive training often required by performance-based direct assessments. This study provides support for both the AR and the BRIEF-P, both rating tools assessing EF in preschoolers. Even though the AR requires training, no additional materials are required for its use, unlike direct assessments of EF. Moreover, the AR is an observational measure, not necessitating the direct manipulation of a stimulus set, which translates to a greater flexibility in its applicability. Where there has been limited coverage of the BRIEF-P in settings other than clinical assessment, this study serves to validate its use in more applied settings, such as a preschool classroom or childcare facility. In sum, after demonstrating a significant relation between the AR and BRIEF-P, it is perhaps most useful to choose a specific measure depending on logistical considerations. For instance, the AR can accompany any direct assessor-child interaction, whereas the BRIEF-P offers a less obtrusive approach referencing a broader time frame of behavior.

IMPLICATIONS FOR POLICY AND PRACTICE

Educators, developmentalists, and policymakers should be informed of the importance of factors such as emotionality and EF for young children, especially those preparing for formal education. Many instances can arise daily, in which children without adequate development in one of these aspects can falter, especially academically and socially (e.g., Carlson and Wang, 2007; Denham et al., 2012c; Herndon et al., 2013). Further, given the plethora of undesirable outcomes associated with low levels of EF and greater negative emotionality in early childhood, it becomes self-evident that the early detection, and addressing, of difficulties in both domains be paramount to promote early social and academic success and school adjustment (Blair, 2002; Denham, 2006; Valiente et al., 2012). Especially because EF are considered to be susceptible to early targeting and interventions (Liew, 2012) and emotional competence can be socialized by preschool classroom teachers (Morris et al., 2013), these results should bolster the ongoing call to arms for curricula and interventions promoting social–emotional learning and EF abilities (Morris et al., 2013; Nix et al., 2013). Further, as this research suggests that both EF and emotionality are related to classroom outcomes, we speculate that the current findings showcase that teachers could find measures potentially useful for predicting positive school outcomes.

LIMITATIONS AND FUTURE DIRECTIONS

A number of issues exist within the current study, some of which could be addressed in future research. The first limitation to the current findings is that given the structure of data collection, the statistical analyses used required that estimated parameters not be bidirectional. Given prior research (e.g., Carlson and Wang, 2007; Ursache et al., 2012; Blankson et al., 2013), there is reason to believe that during early childhood, a bidirectional effect can be found between EF and emotionality. Thus, despite our belief that a bidirectional relation exists between emotional and cognitive development, we chose emotionality to be our exogenous latent construct. Given a larger sample size, structural equation modeling may be suitable for reevaluating our findings allowing for EF to also predict emotionality at T1. Furthermore, having data from a third time point could also allow for the data to be analyzed for additional bidirectional effects through the use of a cross-lagged autoregressive model, for example. Another limitation is that data was not collected from the parents. Having a third source of data could provide stronger validity to our findings and reduce the possibility that our findings are artifacts of the school environment. Additionally, including parental views on their child's EF would provide a more representative portrayal of true EF abilities through the inclusion of another context in which young children spend a considerable amount of their time.

Finally, we provide several ideas for future studies. First, collecting neuropsychological data (e.g., fMRI) could provide additional support to corroborate that portion of the brain responsible for emotionality supplement later development of portions in control of EF. Second, although we found support that emotionality positively predicted later EF, it is possible that these effects differ for younger and older preschoolers. We could not begin such investigation because our sample at T2 consisted of more children considered "older" on the BRIEF-P (4:0–5:11) than "younger" children (2:0–3:11). Given the growth that EF

undergoes just in early childhood, obtaining a more balanced sample with an equal age distribution could be useful to examine whether the current findings are moderated by age. Although our findings support the idea that EF and emotionality are intricately related, we cannot dismiss the possibility of untested confounding variables. Two variables come to mind: temperament and socio-economic status. The temperament literature highlights a construct, termed "effortful control" that helps in bridging the gap between emotion and cognition (for a brief review, see Liew, 2012; see also Carlson and Wang, 2007). It could be that children high in effortful control are able to display more positive emotionality and greater cognitive control (i.e., EF); it is an avenue that could be investigated in future studies. Family socio-economic disadvantage has also shown to have impact on the self-regulatory abilities of children (e.g., Raver, 2012; Raver et al., 2013) and should also be investigated as another potential confound. Last, in light of the current findings, we implore future research to evaluate the role of emotionality wherever relations are found with ER, adopting the one-factor framework of emotion will allow for a more thorough and comprehensive investigation into the vast domain of self-regulation.

CONCLUSION

In sum, prior research has evidenced a consistent interrelation between EF and ER. Conceptualizing ER and emotionality as involving unitary processes, this article is one of the first empirical studies to examine whether a similar interrelation exists between emotionality and EF in a preschool population. We hope that our findings, which indicate that emotionality positively predicts later EF, act as a catalyzing agent in understanding the interconnected development of self-regulatory processes. Additionally, we evidenced the use of both observational measures and standardized rating scales as justifiable means of assessing EF skills in early childhood. The acknowledgment of emotionality, which is easily observable within a preschool classroom yet often uninvestigated in the EF and self-regulation literature, warrants future research regarding the implications of early displays of positive and negative affect.

ACKNOWLEDGMENTS

This research was funded by Institute of Education Sciences grant #R305A110730. We are grateful to the many children, families, and teachers who participated in this study, and the directors of the facilities who so cooperatively worked with us. We also thank Timothy Curby and all graduate, undergraduate, and volunteers working in the Child Development Lab for their unstinting assistance in study organization and data collection.

REFERENCES

Anastopoulos, A. D., Smith, T. F., Garrett, M. E., Morrissey-Kane, E., Schatz, N. K., Sommer, J. L., et al. (2011). Self-regulation of emotion, functional impairment, and comorbidity among children with AD/HD. *J. Atten. Disord.* 15, 583–592. doi: 10.1177/1087054710370567

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* 121, 65–94. doi: 10.1037/0033-2909.121.1.65

Barroso, C., Carrión, G. C., and Roldán, J. L. (2010). "Applying maximum likelihood and PLS on different sample sizes: studies on SERVQUAL model and employee behavior model," in *Handbook of Partial Least Squares*, eds V. Esposito Vinzi, W. Chin, J. Hensler, and H. Wold (Heidelberg: Springer), 427–447.

Bassett, H. H., Denham, S., Wyatt, T. M., and Warren-Khot, H. K. (2012). Refining the Preschool Self-Regulation Assessment for use in preschool classrooms. *Infant Child Dev.* 21, 596–616. doi: 10.1002/icd.1763

Bell, M. A., and Wolfe, C. D. (2004). Emotion and cognition: an intricately bound developmental process. *Child Dev.* 75, 366–370. doi: 10.1111/j.1467-8624.2004.00679.x

Belsky, J., Friedman, S. L., and Hsieh, K. H. (2001). Testing a core emotion-regulation prediction: does early attentional persistence moderate the effect of infant negative emotionality on later development? *Child Dev.* 72, 123–133. doi: 10.1111/1467-8624.00269

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., and Domitrovich, C. E. (2008). Executive functions and school readiness intervention: impact, moderation, and mediation in the Head Start REDI program. *Dev. Psychopathol.* 20, 821–843. doi: 10.1017/S0954579408000394

Blair, C. (2002). School readiness: integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *Am. Psychol.* 57, 111–127. doi: 10.1037/0003-066X.57.2.111

Blair, C., and Diamond, A. (2008). Biological processes in prevention and intervention: the promotion of self-regulation as a means of preventing school failure. *Dev. Psychopathol.* 20, 899–911. doi: 10.1017/S0954579408000436

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Blair, C., Zelazo, P. D., and Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Dev. Neuropsychol.* 28, 561–571. doi: 10.1207/s15326942dn2802_1

Blankson, A. N., O'Brien, M., Leerkes, E. M., Marcovitch, S., Calkins, S. D., and Weaver, J. M. (2013). Developmental dynamics of emotion and cognition processes in preschoolers. *Child Dev.* 84, 346–360. doi: 10.1111/j.1467-8624.2012.01841.x

Bridgett, D. J., Oddi, K. B., Laake, L. M., Murdock, K. W., and Bachmann, M. N. (2013). Integrating and differentiating aspects of self-regulation: effortful control, executive functioning, and links to negative affectivity. *Emotion* 13, 47–63. doi: 10.1037/a0029536

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., and Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Child. Res. Q.* 24, 337–349. doi: 10.1016/j.ecresq.2009.06.001

Brody, G. H., Stoneman, Z., and McCoy, J. K. (1994). Forecasting sibling relationships in early adolescence from child temperaments and family processes in middle childhood. *Child Dev.* 65, 771–784. doi: 10.2307/1131417

Bronstein, P., Ginsburg, G. S., and Herrera, I. S. (2005). Parental predictors of motivational orientation in early adolescence: a longitudinal study. *J. Youth Adolesc.* 34, 559–575. doi: 10.1007/s10964-005-8946-0

Cacioppo, J. T., and Berntson, G. G. (1999). The affect structure: architecture and operating characteristics. *Curr. Dir. Psychol. Sci.* 8, 133–137. doi: 10.1111/1467-8721.00031

Cairns, R. B., and Green, J. A. (1979). "How to assess personality and social patterns: observations or ratings," in *The Analysis of Social Interactions: Methods, Issues, and Illustrations*, ed. R. B. Cairns (Hillsdale, NJ: Erlbaum), 209–226.

Calkins, S. D., and Marcovitch, S. (2010). "Emotion regulation and executive functioning in early development: integrated mechanisms of control supporting adaptive functioning," in *Child Development at the Intersection of Emotion and Cognition*, eds S. D. Calkins and M. A. Bell (Washington, DC: American Psychological Association), 37–57.

Campos, J. J., Frankel, C. B., and Camras, L. (2004). On the nature of emotion regulation. *Child Dev.* 75, 377–394. doi: 10.1111/j.1467-8624.2004.00681.x

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M., and Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cogn. Dev.* 22, 489–510. doi: 10.1016/j.cogdev.2007.08.002

Chaplin, T. M., and Aldao, A. (2013). Gender differences in emotion expression in children: a meta-analytic review. *Psychol. Bull.* 139, 735–765. doi: 10.1037/a0030737

Clark, L. A., and Watson, D. (1995). Constructing validity: basic issues in objective scale development. *Psychol. Assess.* 7, 309–319. doi: 10.1037/1040-3590.7.3.309

Cole, P. M., Martin, S. E., and Dennis, T. A. (2004). Emotion regulation as a scientific construct: methodological challenges and directions for child developmental research. *Child Dev.* 75, 317–333. doi: 10.1111/j.1467-8624.2004.00673.x

Cole, P. M., Zahn-Waxler, C., Fox, N. A., Usher, B. A., and Welsh, J. D. (1996). Individual differences in emotion regulation and behavior problems in preschool children. *J. Abnorm. Psychol.* 105, 518–529. doi: 10.1037/0021-843X.105.4.518

Cowan, P. A., Cohn, D. A., Cowan, C. P., and Pearson, J. L. (1996). Parents' attachment histories and children's externalizing and internalizing behaviors: exploring family systems models of linkage. *J. Consult. Clin. Psychol.* 64, 53–63. doi: 10.1037/0022-006X.64.1.53

Davidson, R. J., Jackson, D. C., and Kalin, N. H. (2000). Emotion, plasticity, context, and regulation: perspectives from affective neuroscience. *Psychol. Bull.* 126, 890–909. doi: 10.1037/0033-2909.126.6.890

Davies, P. T., and Cummings, E. M. (1998). Exploring children's emotional security as a mediator of the link between marital relations and child adjustment. *Child Dev.* 69, 124–139. doi: 10.2307/1132075

Denham, S. A. (2006). Social-emotional competence as support for school readiness: what is it and how do we assess it? *Early Educ. Dev.* 17, 57–89. doi: 10.1207/s15566935eed1701_4

Denham, S. A. (2007). Dealing with feelings: how children negotiate the worlds of emotions and social relationships. *Cogn. Brain Behav.* 11, 1–48.

Denham, S. A., Warren-Khot, H. K., Bassett, H. H., Wyatt, T., and Perna, A. (2012a). Factor structure of self-regulation in preschoolers: testing models of a field-based assessment for predicting early school readiness. *J. Exp. Child Psychol.* 111, 386–404. doi: 10.1016/j.jecp.2011.10.002

Denham, S. A., Bassett, H. H., Thayer, S. K., Mincic, M., Sirotkin, Y. S., and Zinsser, K. (2012b). Observing preschoolers' social-emotional behavior: structure, foundations, and prediction of early school success. *J. Genet. Psychol.* 173, 246–278. doi: 10.1080/00221325.2011.597457

Denham, S. A., Bassett, H. H., Mincic, M. M., Kalb, S. C., Way, E., Wyatt, T., et al. (2012c). Social-emotional learning profiles of preschoolers' early school success: a person-centered approach. *Learn. Individ. Differ.* 22, 178–189. doi: 10.1016/j.lindif.2011.05.001

Denham, S. A., Bassett, H. H., Sirotkin, Y. S., and Zinsser, K. (2013). Head Start preschoolers' emotional positivity and emotion regulation predict their social-emotion behavior, classroom adjustment, and early school success. *NHSA Dialog* 16, 18–33.

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach-Major, S., et al. (2003). Preschool emotional competence: pathway to social competence. *Child Dev.* 74, 238–256. doi: 10.1111/1467-8624.00533

Denham, S. A., Caverly, S., Schmidt, M., Blair, K., DeMulder, E., Caal, S., et al. (2002). Preschool understanding of emotions: contributions to classroom anger and aggression. *J. Child Psychol. Psychiatry* 43, 901–916. doi: 10.1111/1469-7610.00139

Duncan, G. J. (2012). Give us this day our daily breadth. *Child Dev.* 83, 6–15. doi: 10.1111/j.1467-8624.2011.01679.x

Efron, B., and Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *Am. Stat.* 37, 36–48. doi: 10.1080/00031305.1983.10483087

Eisenberg, N., Fabes, R. A., Guthrie, I. K., and Reiser, M. (2000). Dispositional emotional and regulation: their role in predicting quality of social functioning. *J. Pers. Soc. Psychol.* 78, 136–157. doi: 10.1037/0022-3514.78.1.136

Falk, R. F., and Miller, N. B. (1982). *A Primer for Soft Modeling*. Akron: The University of Akron Press.

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50. doi: 10.2307/3151312

Garon, N., Bryson, S. E., and Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi: 10.1037/0033-2909.134.1.31

Gioia, G. A., Espy, K. A., and Isquith, P. K. (2003). *Behavior Rating Inventory of Executive Function-Preschool Version*. Lutz, FL: Psychological Assessment Resources.

Graziano, P. A., Calkins, S. D., and Keane, S. P. (2011). Sustained attention development during the toddlerhood to preschool period: associations with toddlers' emotion regulation strategies and maternal behaviour. *Infant Child Dev.* 20, 389–408. doi: 10.1002/icd.731

Henseler, J., Ringle, C. M., and Sinkovics, R. R. (2009). "The use of partial least squares path modeling in international marketing," in *Advances in International Marketing*, eds R. R. Sinkovics and P. N. Ghauri (Bingley, UK: Emerald Group Publishing Limited), 277–319.

Herndon, K. J., Bailey, C. S., Shewark, E., Denham, S. A., and Bassett, H. H. (2013). Preschoolers' emotion expression and regulation: relations with school adjustment. *J. Genet. Psychol.* 174, 642–663. doi: 10.1080/00221325.2012.759525

Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Dev. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Iida, S., Nakao, T., and Ohira, H. (2011). Implicit attenuation of subsequent emotion by cognitive activity. *Cogn. Affect. Behav. Neurosci.* 11, 476–484. doi: 10.3758/s13415-011-0045-y

Isley, S. L., O'Neil, R., Clatfelter, D., and Parke, R. D. (1999). Parent and child expressed affect and children's social competence: modeling direct and indirect pathways. *Dev. Psychol.* 35, 547–560. doi: 10.1037/0012-1649.35.2.547

Isquith, P. K., Gioia, G. A., and Espy, K. A. (2004). Executive function in preschool children: examination through everyday behavior. *Dev. Neuropsychol.* 26, 403–422. doi: 10.1207/s15326942dn2601_3

Izard, C. E. (2002). Translating emotion theory and research into preventive interventions. *Psychol. Bull.* 128, 796–824. doi: 10.1037/0033-2909.128.5.796

Jahromi, L. B., and Stifter, C. A. (2008). Individual differences in preschoolers' self-regulation and theory of mind. *Merrill Palmer Q.* 54, 125–150. doi: 10.1353/mpq.2008.0007

Kanske, P., and Kotz, S. A. (2012). Effortful control, depression, and anxiety correlate with the influence of emotion on executive attentional control. *Biol. Psychol.* 91, 88–95. doi: 10.1016/j.biopsycho.2012.04.007

Lehtonen, A., Howie, E., Trump, D., and Hudson, S. M. (2012). Behavior in children with neurofibromatosis type 1: cognition, executive function, attention, emotion, and social competence. *Dev. Med. Child Neurol.* 55, 111–125. doi: 10.1111/j.1469-8749.2012.04399.x

Liew, J. (2012). Effortful control, executive functions, and education: bringing self-regulatory and social-emotional competencies to the table. *Child Dev. Perspect.* 6, 105–111. doi: 10.1111/j.1750-8606.2011.00196.x

Luu, P., Collins, P., and Tucker, D. M. (2000). Mood, personality, and self-monitoring: negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *J. Exp. Psychol.* 129, 43–60. doi: 10.1037/0096-3445.129.1.43

Marjoribanks, K. (1997). Family background, social and academic capital, and adolescents' aspirations: a mediational analysis. *Soc. Psychol. Educ.* 2, 177–197. doi: 10.1023/A:1009602307141

Martel, M. M. (2009). Research review: a new perspective on attention-deficit/hyperactivity disorder: emotion dysregulation and trait models. *J. Child Psychol. Psychiatry* 50, 1042–1051. doi: 10.1111/j.1469-7610.2009.02105.x

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., and Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Dev. Psychol.* 43, 947–959. doi: 10.1037/0012-1649.43.4.947

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., and Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Morris, C. A., Denham, S. A., Bassett, H. H., and Curby, T. W. (2013). Relations among teachers' emotion socialization beliefs and practices and preschoolers' emotional competence. *Early Educ. Dev.* 24, 979–999. doi: 10.1080/10409289.2013.825186

Morris, P., Millenky, M., Raver, C. C., and Jones, S. M. (2013). Does a preschool social and emotional learning intervention pay off for classroom instruction and children's behavior and academic skills? Evidence from the Foundations of

Learning Project. *Early Educ. Dev.* 24, 1020–1042. doi: 10.1080/10409289.2013.825187

Nelson, C. A. (1994). "Neural bases of infant temperament," in *Temperament: Individual Differences at the Interface of Biology and Behavior*, eds J. Bates and T. E. Wachs (Washington, DC: American Psychological Association), 47–82. doi: 10.1037/10149-002

Nigg, J. T., Martel, M. M., Nikolas, M., and Casey, B. J. (2010). "Intersection of emotion and cognition in developmental psychopathology," in *Child Development: At the Intersection of Emotion and Cognition*, eds S. D. Calkins and M. A. Bell (Washington, DC: APA Press), 225–246. doi: 10.1037/12059-012

Nix, R. L., Bierman, K. L., Domitrovitch, C. E., and Gill, S. (2013). Promoting children's social-emotional skills in preschool can enhance academic and behavioral functioning in kindergarten: findings from Head Start REDI. *Early Educ. Dev.* 24, 1000–1019. doi: 10.1080/10409289.2013.825565

Raver, C. C. (2012). Low-income children's self-regulation in the classroom: scientific inquiry for social change. *Am. Psychol.* 67, 681–689. doi: 10.1037/a0030085

Raver, C. C., Blair, C., Willoughby, M., and Family Life Project Key Investigators. (2013). Poverty as a predictor of 4-year-olds' executive function: new perspectives on models of differential susceptibility. *Dev. Psychol.* 49, 292–304. doi: 10.1037/a0028343

Riggs, N. R., Jahromi, L. B., Razza, R. P., Dilworth-Bart, J. E., and Mueller, U. (2006). Executive function and the promotion of social-emotional competence. *J. Appl. Dev. Psychol.* 26, 300–309. doi: 10.1016/j.appdev.2006.04.002

Ringle, C. M., Wende, S., and Will, A. (2005). *SmartPLS Release: 2.0 (beta)*. Available at: http://www.smartpls.de

Roid, G. H., and Miller, L. J. (1997). *Social Emotional Rating Scale – Examiner Version. Leiter International Performance Scale – Revised (Leiter-R)*. Wood Dale, IL: Stoelting Co.

Smith-Donald, R., Raver, C. C., Hayes, T., and Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-Regulation Assessment (PSRA) for field-based study. *Early Child. Res. Q.* 22, 173–187. doi: 10.1016/j.ecresq.2007.01.002

Spiliotopoulou, G. (2009). Reliability reconsidered: Cronbach's alpha and paediatric assessment in occupational therapy. *Aust. Occup. Ther. J.* 56, 150–155. doi: 10.1111/j.1440-1630.2009.00785.x

Thompson, R. A., Lewis, M. D., and Calkins, S. D. (2008). Reassessing emotion regulation. *Child Dev. Perspect.* 2, 124–131. doi: 10.1111/j.1750-8606.2008.00054.x

Trentacosta, C. J., and Izard, C. E. (2007). Kindergarten children's emotion competence as a predictor of their academic competence in first grade. *Emotion* 7, 77–88. doi: 10.1037/1528-3542.7.1.77

Tsethlikai, M. (2010). The influence of a friend's perspective on American Indian children's recall of previously misconstrued events. *Dev. Psychol.* 46, 1481–1496. doi: 10.1037/a0020725

Tsethlikai, M. (2011). An exploratory analysis of American Indian children's cultural engagement, fluid cognitive skills, and standardized verbal IQ scores. *Dev. Psychol.* 47, 192–202. doi: 10.1037/a0020803

Ursache, A., Blair, C., and Raver, C. C. (2012). The promotion of self-regulation as a means of enhancing school readiness and early achievement in children at risk for school failure. *Child Dev. Perspect.* 6, 122–128. doi: 10.1111/j.1750-8606.2011.00209.x

Ursache, A., Blair, C., Stifter, C., Voegtline, K., and The Family Life Project Investigators. (2013). Emotional reactivity and regulation in infancy interact to predict executive functioning in early childhood. *Dev. Psychol.* 49, 127–137. doi: 10.1037/a0027728

Valiente, C., Swanson, J., and Eisenberg, N. (2012). Linking students' emotions and academic achievement: when and why emotions matter. *Child Dev. Perspect.* 6, 129–135. doi: 10.1111/j.1750-8606.2011.00192.x

Wakschlag, L. S., Leventhal, B. L., Briggs-Gowan, M. J., Danis, B., Keenan, K., Hill, C., et al. (2005). Defining the "disruptive" in preschool behavior: what diagnostic observation can teach us. *Clin. Child Fam. Psychol. Rev.* 8, 183–201. doi: 10.1007/s10567-005-6664-5

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 13 January 2014; accepted: 05 May 2014; published online: 27 May 2014.
*Citation: Ferrier DE, Bassett HH and Denham SA (2014) Relations between executive function and emotionality in preschoolers: Exploring a transitive cognition–emotion linkage. Front. Psychol. **5**:487. doi: 10.3389/fpsyg.2014.00487*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Ferrier, Bassett and Denham. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 25 August 2014
doi: 10.3389/fpsyg.2014.00922

Child temperamental reactivity and self-regulation effects on attentional biases

Georgiana Susa[1], Oana Benga[1]*, Irina Pitica[1] and Mircea Miclea[2]

[1] Developmental Psychology Lab, Department of Psychology, Babes-Bolyai University, Cluj-Napoca, Romania
[2] Department of Psychology, Babes-Bolyai University, Cluj-Napoca, Romania

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Hwajin Yang, Singaopre Management University, Singapore
IKUKO Shinohara, Aichi Shukutoku University, Japan

***Correspondence:**
Oana Benga,
Developmental Psychology Lab,
Department of Psychology,
Babes-Bolyai University, Republicii Str., No 37, Cluj-Napoca 400015, Romania
e-mail: oanabenga@gmail.com

This study examined the effects of individual differences in temperamental reactivity (fear) and self-regulation (attentional control) on attentional biases toward threat in a sample of school-aged children (age range was between 9 years 1 month and 13 years 10 months). Attentional biases were assessed with pictorial Dot-probe task, comparing attention allocation toward angry (threat-related) vs. neutral and happy faces. Children also completed self-report temperamental measures of fear and attentional control. We compared attentional bias scores in 4 groups of children: high/low fear and high/low attentional control. Results indicated that, in the case of children with high fear and low attentional control, attention was significantly biased toward angry faces compared with children who had low fear and low attentional control. Findings are discussed in terms of the moderating role of individual differences in attentional control in the context of threat, anxiety-related attentional biases in children.

Keywords: child temperament, fear, attentional control, attentional biases, anxiety

INTRODUCTION

Cognitive theories have proposed that anxious individuals tend to direct their attention toward threatening information during early stages of processing (Beck and Clark, 1997). Specific theoretical accounts of attentional biases toward threat state that biases could appear as a result of exaggerated pre-attentional threat evaluation (Williams et al., 1988; Mogg and Bradley, 1998), but also as a result of a failure of effortful strategies to focus on task-related rather than threat stimuli (Mathews and MacKintosh, 1998; Cisler and Koster, 2010).

A lot of research in this field has been conducted with adults and indicates that the tendency to attend toward threatening information is associated with both anxiety disorders and non-clinical high levels of anxiety (Bar-Haim et al., 2007). Childhood investigations have also examined the association between attentional biases toward threat and anxiety, with different experimental paradigms, from reaction time (e.g., Dot-probe and emotional Stroop tasks) to eye-tracking (see In-Albon et al., 2010), but results are less straightforward.

Attentional biases were found in clinical child populations, when using verbal (neutral vs. threat words) or non-verbal stimuli (negative, e.g., angry or fearful vs. neutral and positive facial expressions) (e.g., Vasey et al., 1995; Dalgleish et al., 2003; Brotman et al., 2007; Roy et al., 2008; Waters et al., 2010a). Moreover, in non-clinical samples, some studies comparing children categorized as highly anxious vs. non-anxious reported attentional biases characteristic of the first category, when using both the emotional Stroop task (Martin and Jones, 1995; Richards et al., 2000) and the Dot-probe task (Vasey et al., 1996). A recent study compared low-anxious and moderately-anxious children aged 7–14 years in an emotional Stroop task as well as in a face Dot-probe task. Authors reported enhanced processing bias for angry faces compared to neutral ones on the Stroop task as being characteristic of younger moderately anxious children, mean age 9 (Reinholdt-Dunne et al., 2012). The same children showed enhanced negative as well as positive processing biases in the Dot-probe task. However, older moderately anxious children (mean age 11) had lessened anxiety-related threat bias, result interpreted as a consequence of heightened abilities of executive control which are growing with age.

In spite of the above-mentioned evidence for threat-related attentional biases in anxious children, several studies reported challenging results. For example, whereas studies generally found evidence for attentional biases toward threat (e.g., Roy et al., 2008), some authors reported a pattern of attentional biases away from threatening stimuli (e.g., Monk et al., 2006). Furthermore, some studies revealed that attentional biases are present in both anxious and non-anxious children (e.g., Eschenbeck et al., 2004), whereas others found attentional biases toward threat only in children who were clinically diagnosed with anxiety disorders (Roy et al., 2008). Also, results with regard to attentional biases for positive stimuli (happy faces) are mixed and rather ambiguous (e.g., Waters et al., 2008, 2010b; Reinholdt-Dunne et al., 2012).

In conclusion, there seem to be diverging, thus provocative findings on threat-related attentional biases in children. In addition, an essential question addresses the role of these attentional biases on the onset and maintenance of clinical anxiety. Both lines of inquiry recently encouraged the growth of developmental perspectives that look at potential routes through which such biases toward threat might emerge in children and further sustain the development of psychopathology.

TEMPERAMENT AND ATTENTIONAL BIASES

A fundamental theoretical position suggests that temperamental factors might predispose children to manifest attentional biases toward threat (Lonigan et al., 2004; Helzer et al., 2009; Pine et al., 2009).

According to the model of Lonigan et al. (2004), temperamental traits involving sensitivity toward threat (e.g., negative affectivity, behavioral inhibition) are considered critical in making children prone to allocate their attention toward threatening information. Negative affectivity represents a reactive dimension of temperament, within the theoretical framework of Rothbart (Rothbart and Derryberry, 1981), and is defined as a tendency to experience negative emotions, such as fear, sadness, and anger. This reactive dimension of temperament describes the individual's responsiveness to environmental stimuli, in terms of the extent to which negative affect and avoidance behaviors vs. positive affect and approach behaviors are elicited (Henderson and Wachs, 2007; Posner and Rothbart, 2007). Also, this temperamental dimension is characterized by sensitivity to negative stimuli, physiological arousal, and emotional distress (Ingram and Price, 2010). A different approach, first advanced by Kagan (Kagan et al., 1987; Schmidt et al., 1997; Fox et al., 2005) delineates the temperamental profile of behavior inhibition, characterized by fear of novelty, social reticence, and further proneness to internalizing disorders, thus partly overlapping with a high negative affectivity temperamental profile.

Along with such temperamental traits that might underlie an individual's sensitivity to threat, the model of Lonigan et al. (2004) takes into consideration the self-regulatory system, serving to modulate reactivity, which is described in Rothbart's theory by the dimension of effortful control. This regulative dimension of temperament includes processes such as inhibition, avoidance, and attentional self-regulation (Rothbart and Bates, 2006; Posner and Rothbart, 2007).

A specific component of the self-regulatory temperamental dimension of effortful control is that of attentional control. Temperamental attentional control reflects stable individual differences in the ability to focus and shift attention with ease vs. harder or less natural (Posner and Rothbart, 2000; Simonds et al., 2007), assumed to have its cognitive underpinnings in the executive or anterior attention system (Posner and DiGirolamo, 1998; further developments in Fan et al., 2002; Rueda et al., 2005); but see also a recent approach advanced by Zhou et al. (2012) that considers temperamental attentional control similar to executive control. A critical function of attentional control is to disengage attention from threatening irrelevant information and to keep attention focused on task relevant stimuli. Thus, attentional control is considered especially important in emotion regulation in general and in reducing internalizing symptoms such as anxiety and sadness in particular (Eisenberg et al., 2000; Calkins and Fox, 2002; Bell and Calkins, 2012).

The model advanced by Lonigan and collaborators suggests that attentional biases for threat can be seen in children with high levels of negative affectivity who also have low effortful control and more specifically low attentional control. From a developmental perspective, attentional biases are expected to emerge early in life in children born with an underlying anxiety predisposition, such as high levels of negative affectivity, coupled with low levels of effortful control. Moreover, such biases are expected to further play a mediating role in the relation between temperament and the development of anxiety disorders. Although this theoretical model is compelling, to date there are few studies that specifically examined the link between temperament and attentional biases toward threat. Initial results are nevertheless promising, suggesting that children with fearful temperament—an important aspect of negative affectivity—tend to preferentially allocate their attention toward threat (White et al., 2010). Moreover, children with high levels of both negative affectivity and attentional control do not show attention biases toward threatening words, while children with high levels of negative affectivity coupled with low attentional control present vigilance toward these stimuli, as demonstrated by Lonigan and Vasey (2009) with a Dot-probe task using neutral and threatening words. Efficient attentional control processes may help children with fearful temperament inhibit the processing of task-irrelevant information and focus on the task-relevant information in the environment (see Pine et al., 2009 for a theoretical interpretation of data). High attentional control can thus enable individuals to override initial reactive attentional biases, and further serve as a protective factor against the development of anxiety disorders, as demonstrated by Vervoort et al. (2011) in a study with adolescents. In addition, empirical research on anxiety and attention, in adults with high levels of trait anxiety, has provided evidence that impaired attentional control might underlie attentional biases. Individuals with high trait anxiety but low levels of self-reported attentional control maintained a vigilance bias toward threat cues even at 500 ms, whereas those with high levels of attentional control shifted attention away from the threat location (Derryberry and Reed, 2002). From a developmental perspective, the longitudinal study conducted by Hardee et al. (2013) showed, in an event-related functional magnetic resonance imaging with Dot-probe task, that young adults characterized in early childhood with behavioral inhibition (BI) exhibited greater strength in threat-related connectivity than non-behaviorally inhibited young adults. Specifically, young adults with a history of BI manifested greater negative connectivity between amygdala and two frontal regions (dorsolateral prefrontal cortex and anterior insula) during trials containing angry faces compared to neutral faces. Also, amygdala—insula connectivity interacted with childhood BI to predict young adult internalizing symptoms.

All these results highlight the importance of analyzing the role of regulative temperamental factors, such as attentional control processes, in junction with the role of reactive temperamental traits, like negative affectivity or fear. Also, these results converge with cognitive theoretical accounts of attentional biases (Mogg and Bradley, 1998; Cisler and Koster, 2010), in suggesting that exaggerated engagement of attention to threat is, on the one hand, linked with an automatic/ pre-attentional threat detection mechanism, which is extremely sensitive in people born with an underlying anxiety predisposition, and on the other hand, with a failure of effortful strategies such as temperamental attentional control to regulate these initial automatic tendencies. However, in children, the relation between temperamental variables (both reactive and regulative) and attention toward threat has been

under-investigated (with the exception of the studies done by Helzer et al., 2009; Lonigan and Vasey, 2009). Moreover, childhood research discussed above used threat-related words rather than emotion-eliciting pictorial stimuli. But pictorial stimuli, for example emotional facial expressions, are considered more ecological, compared to words which are limited in threat value and more open to subjective interpretability (Mogg and Bradley, 1999).

PRESENT STUDY

In the present study, we aimed to examine the effects of individual differences in temperamental fear and attentional control processes on attention allocation toward threat, in children aged 9–14. Based on the assumed strengths of the Lonigan model, of greatest interest was the interaction effect between temperamental fear and temperamental attentional control on attentional allocation toward threatening information. We examined attentional biases toward threatening facial expressions, in order to fill the gap in existing research relative to attentional processing of more ecological stimuli.

Our hypotheses were the following: first, regarding the influence of temperamental fear, we expected that children with higher levels of fear would show enhanced attentional allocation toward angry faces, compared to children with low fear; second, we expected that temperamental attentional control might moderate the relation between temperamental fear and threat-related attentional biases. Specifically, only fearful children with low attentional control were expected to significantly bias their attention toward angry faces.

We believe that, from a theoretical perspective, the present study will extend the existing research on the relation between temperamental predispositions and threat related attentional biases, by adding information about the mechanisms underlying the emergence of attentional biases. Moreover, such an approach may help to inform prevention strategies regarding children who are prone to develop anxiety disorders. Such strategies could be designed to increase their resilience by means of attentional control training procedures.

MATERIALS AND METHODS

PARTICIPANTS

Our sample initially consisted of 163 school-aged Romanian children. This sample was part of a larger screening study conducted in our laboratory, concerning the relations between attentional biases and anxiety symptoms in children and adolescents. We obtained parental written informed consent and verbal consent from each child before the testing. In the current study, we included only children for whom we had both reaction time data and self-report data for temperamental fear, temperamental attentional control, and non-clinical anxiety symptoms. Consequently, 5 children were excluded from this study due to scheduling difficulties that lead to missing data on the measure of anxiety symptoms. The final sample included 158 participants, 70 of them girls. The age range of these participants was between 9 years 1 month and 13 years 10 months. Mean age of this sample was 11 years and 3 months. All children included in the sample were free of any clinical psychological diagnosis, as reported by teachers and school psychologists. Also their vision was normal or corrected.

MATERIALS

The questionnaires employed in this study to assess temperamental variables were the fear subscale from the Early Adolescent Temperament Questionnaire-Revised (EATQ-R; Ellis and Rothbart, 2001) and the child version of the Attentional Control Scale (ACS-C; Derryberry and Reed, 2002). Even though EATQ-R assesses various components of temperament, we selected only the fear subscale, since our research question is grounded on previous data showing that children with temperamental fear– an important aspect of negative affectivity—tend to preferentially allocate their attention toward threat (White et al., 2010). EATQ-R also contains an attentional control subscale, but this has only 8 items compared to ACS-C that has 20, thus being a more comprehensive measure of this temperamental dimension. The rationale for choosing these two particular temperamental scales was that both were developed based on Rothbart's model of temperament, which represents the conceptual temperamental framework of the Lonigan model.

The EATQ-R is a measure of temperament designed to be used with 9 to 15 year old children and adolescents. We selected the fear subscale of this questionnaire to assess self-reported temperamental fear in children. The fear subscale reflects the tendency to experience unpleasant anticipation of distress (Helzer et al., 2009). Children are asked to rate each item on a 5-point Likert scale and assess the frequency with which the item is true or false in their case. Some examples of items from the fear subscale of the EATQ-R are: "I worry about getting into trouble" or "I worry about my parent(s) dying or leaving me." The EATQ-R was adapted for use with Romanian children through the following steps: (a) the scale was translated from English into Romanian by an expert in the field of temperament and development; (b) in order to verify that the original conceptual content has been preserved in the Romanian version, the Romanian translation was back translated to English by a different expert with proficiency level English as a foreign language qualifications; (c) the Romanian translation of the EATQ-R was employed in a pilot study with children aged between 9 and 14, to check that the language used was accessible to this age group.

In the present study we used only the fear subscale of EATQ-R that showed moderate internal consistency, Cronbach's Alpha being 0.69 in our sample of children.

The ACS-C is a self-report 20-item scale that evaluates children's ability to focus and shift attention. The scale contains 10 items that measure the ability to focus attention (e.g., "When I concentrate myself, I do not notice what is happening in the room around me") and another 10 items that measure the ability to shift attention (e.g., "When I am doing something, I can easily stop and switch to some other task"). Children are answering the items by reporting how frequently certain things happen to them and they respond on a 4-point Likert scale. A good capacity of attentional control is reflected by higher scores obtained on this scale. Different studies conducted with different samples report good internal consistency of the ACS-C (Muris et al., 2004, 2007). The ACS-C was adapted for use with Romanian children through the

same procedure described in the case of the EATQ-R adaptation. In the present study the ACS-C showed good internal consistency as Cronbach's Alpha coefficient reached 0.80.

The Spence Child Anxiety Scale was used to measure anxiety symptoms (SCAS; Spence, 1998). The SCAS child version is a 38-item self-report anxiety measure. This questionnaire asks children to rate how frequently they experience the situations described by each item using a 4-point Likert scale: 1- Never, 2- Sometimes, 3- Often, and 4- Always. By summing the scores from all items a total score can be computed. Also the SCAS offers subscale scores based on the anxiety disorder categories indexed in the Diagnostic and Statistical Manual for Mental Disorders IV (American Psychiatric Association, 1994). The subscales assess social anxiety, separation anxiety, obsessive-compulsive disorder, panic and agoraphobia, physical injury fears, and generalized anxiety. The Romanian version of the SCAS has been adapted for use with Romanian children through the same procedure described in the case of the EATQ-R adaptation and is currently under validation (Benga et al., in press). Studies conducted with other samples reported good psychometric properties (Spence, 1998; Spence et al., 2003). In the current study we obtained good internal consistency for the global scale. Cronbach's Alpha coefficient reached 0.85.

Attentional biases were measured with a pictorial version of the Dot-probe task, adapted from Bradley et al. (1998) and Susa et al. (2012). During the task, the children were seeing a series of trials on the computer screen. Each trial consisted of the following events: the fixation point in the center of the screen for 500 ms, a pair of pictures showing human facial expressions for 500 ms, the probe (a star) replacing one of the pictures, and a blank screen as a pause for 500 ms. The probe was displayed on the screen until a response was given. The facial stimuli were 64 images selected from a pool of 96 images from the following sets: 22 from the NimStim (Tottenham et al., 2009; http://www.macbrain.org/resources.htm)[1], 5 from the Ekman stimuli set (Ekman and Friesen, 1976) and 37 from the stimuli developed by Mogg and Bradley (Bradley et al., 1998). We combined stimuli from different sets in order to present only Caucasian persons, since Romanian children are mostly familiar with this race. In the current study, we did not ask children to complete an emotion recognition task due to time constraints. However, we recruited a second group of children having, the same age as participants from our initial sample, and we tested whether they can accurately identify the emotional meaning (i.e., recognition accuracy) and rate the emotional intensity of the facial stimuli used within the Dot-Probe task (the description of this study and its results can be found in the Supplemental Material)[2]. To our knowledge, there are no other published validation studies with children for these picture sets. Though, face stimuli from the three databases were in general used by all previous studies conducted with children (e.g., for Ekman stimuli set see Szpunar and Young, 2012; for Mogg and Bradley stimuli see Roy et al., 2008; for NimStim see Tottenham et al., 2011) and data seem to support the view that children can recognize the emotional meaning at adult-like levels.

Two types of threatening facial expressions were used in previous studies, in order to assess attentional allocation toward emotional facial expressions, namely fearful and angry faces. To our knowledge, there are no studies reported in the literature that compared attentional biases to fearful and angry faces in anxious children. However, a study conducted by Mogg et al. (2007) with an adult sample showed that fearful and angry faces elicited similar attentional biases in high-anxious individuals. In general, fearful faces were used by neuroimaging research (e.g., Whalen et al., 2001) since they seem to elicit more amygdala activity, given that they are more ambiguous (e.g., they signify the presence of danger, but do not provide information about its source). In contrast, angry faces were predominantly employed by Dot-probe studies, in which research questions were framed in terms of cognitive models of anxiety and which investigated the influence of anxiety on attentional allocation toward threat. Therefore, to facilitate the comparison and interpretation of our data with previous Dot-probe studies, we chose to present angry faces in order to assess attentional biases for threat.

PROCEDURE

Data from both the questionnaires and the Dot-probe task were collected from children in two schools, in the presence of a research assistant. Children who voluntarily consented to participate were asked to have their parents sign the informed consent form. In order to prevent children's fatigue, questionnaire data were collected first and then, approximately 2 weeks after, children completed the Dot-probe task.

For the Dot-probe task, children were seated in front of the computer at a distance of approximately 40 cm from the screen. At this distance, they were able to comfortably reach the laptop keyboard throughout the task. The task was presented to the children as a computer game and they were asked to read the instructions displayed. Participants were instructed to press key A when the probe replaced the picture on the left side of the screen and key L when the probe replaced the picture on the right side of the screen (on a QWERTY keyboard). Before starting the task, the research assistant summarized for each child what he or she was asked to do. For each child, the program presented the picture pairs in random order. At the end, each child received positive feedback and a small reward.

During the Dot-probe task all children included in the analysis completed 160 experimental trials and 8 practice trials. There were a total of 80 pairs of stimuli, 32 of them showing angry-neutral facial expressions, 32 showing happy-neutral facial expressions and 16 pairs showing neutral- neutral facial expressions. By including neutral-neutral pairs, we could analyze the two mechanisms of attentional biases discussed in the literature: attentional faster engagement by angry faces or difficulty of disengagement from angry faces (e.g., Koster et al., 2004). Also, in this way we could control that our reaction time data are not better explained by behavioral interference effects (Wolters et al., 2012).

RESULTS

PRELIMINARY ANALYSES

Dot-probe data preparation

Reaction time data for each participant were screened and trials with response time less than 200 ms or greater than 1500 ms were

[1]The MacArthur Foundation Research Network on Early Experience and Brain Development. (n.d.).

[2]We are grateful to Reviewer 1 for suggesting this additional data collection in order to validate the emotional face stimuli used in the Dot-Probe task.

eliminated from further analyses (a total of 0.67% of the total data). Trials with reaction times greater than 1500 ms are considered to represent outliers and are likely attributable to error, therefore not excluding these reaction times would have influenced individually trimmed means (Oehlberg et al., 2012). It is highly probable that during such trials children were not paying close attention to the displayed stimuli. Mean accuracy level for the whole sample was 98.73% of all responses. Trials with incorrect responses were excluded from the reaction time analysis.

We then computed an attentional bias score for each child. These bias scores were calculated by subtracting mean reaction times for congruent trials from mean reaction times for incongruent trials (Mogg and Bradley, 1999). The difference between congruent and incongruent trials is the location of the probe relative to the emotional face. In congruent trials, the probe appeared on the same location as the emotional face (angry or happy), and for incongruent trials the probe appeared on the same location as the neutral face. Positive values indicate a vigilance bias and negative values indicate an avoidance bias for emotional faces. The same analysis was carried out both for angry as well as for happy expression trials.

Questionnaire total sample and group characteristics

The total mean fear score on the fear subscale of the EATQ, for the whole sample of children in this study was 2.85 ($SD = 0.75$; minimum score 1, maximum score 4.67). This is similar to the mean reported by Muris and Meesters (2009) in a community sample of Belgian and Dutch children aged 8 to 14 ($M = 2.69$, $SD = 0.77$).

The total mean attentional control score on the ACS-C for the whole sample of children in this study was 26.74 ($SD = 6.14$; minimum score 11, maximum score 44). This is somewhat different from the mean reported by Muris et al. (2004) in a community sample of Dutch children aged 8 to 13 ($M = 34$, $SD = 8.1$).

Mean anxiety score on the SCAS, in the whole sample, was 29.28 ($SD = 15.88$; minimum score 1, maximum score 81). The Romanian version of the SCAS is currently under validation but preliminary data (Benga et al., in press) from a sample of 300 children aged between 9 and 15 years showed a similar SCAS mean score ($M = 29.60$, $SD = 15.43$; minimum score 1, maximum score 82).

By using the median split of the ratings of children's fear level (median value was 3) and attentional control level (median value was 20) we formed four groups (see **Table 1** for descriptive data within each condition): a high fear, high attentional control group (HFHAC); a high fear, low attentional control group (HFLAC); a low fear, high attentional control group (LFHAC); and a low fear, low attentional control group (LFLAC). The four groups did not significantly differ in age, $F_{(3,\ 154)} = 0.87$, *ns* or anxiety scores, $F_{(3,\ 154)} = 1.10$, *ns*. Also, the four groups did not significantly differ in overall reaction times in the Dot-probe task, $F_{(3,\ 154)} = 0.49$, *ns*, or in accuracy, $F_{(3,\ 154)} = 0.27$, *ns*.

In order to control for behavioral interference effects in the Dot-probe task, we conducted a preliminary 2 × 2 × 2 ANCOVA (Fear x Attentional control × Face valence) analysis with Age as a covariate, for comparing reaction times in the four groups between all conditions, with neutral faces collapsed and all conditions with angry faces collapsed. We ran separately a 2 × 2 × 2 ANCOVA analysis between all conditions with neutral faces and all conditions with happy faces. Results indicated no significant differences in overall reaction times when face stimuli were neutral vs. angry, $F_{(5,\ 152)} = 0.23$, *ns*, or when face stimuli were neutral vs. happy, $F_{(5,\ 152)} = 1.99$, *ns*.

MAIN ANALYSES

The theoretical focus of our study was on estimating the impact of temperamental variables and of their interaction on attentional biases, while controlling for the effect of other variables that may influence both the measured independent and the dependent variables. Therefore, because anxiety may influence attention to threat and individual differences in temperamental traits are associated with anxiety, we included anxiety as a covariate. Also, the quasi-experimental design of the present study, in which participants were not randomly assigned to groups, requires the inclusion of this covariate (Yzerbyt et al., 2004). Besides anxiety, we also included age as a covariate in the design, since our sample covered quite a wide age range and this is a factor known to influence reaction times (Anderson et al., 1997; Iida et al., 2010).

Analysis of covariance

As our hypotheses were concerned with differences between groups, we conducted a mixed ANCOVA with Emotion valence (angry or happy) as a within-subjects factor, Fear and Attentional Control levels as between-subjects factors, and Age (in months)

Table 1 | Descriptive data for each group as a function of both temperamental dimension (fear and attentional control), gender and age.

Group	*N*	Gender		Age			Fear		Attentional control		Age in months	
		Girls	Boys	9–10	11–12	13–14	Mean	*SD*	Mean	*SD*	Mean	*SD*
HFHAC[a]	23	11	12	9	9	5	3.55	0.29	23.39	2.13	133.39	15.20
		47.8%	52.2%	39.13%	39.13%	21.74%						
HFLAC[b]	43	21	22	11	25	7	3.63	0.39	15.60	2.92	136.30	12.69
		48.84%	51.16%	25.58%	58.14%	16.28%						
LFHAC[c]	53	21	32	15	21	17	2.35	0.54	24.28	2.09	137.98	15.40
		39.6%	60.4%	28.31%	39.62%	32.07%						
LFLAC[d]	39	17	22	11	19	9	2.39	0.49	16.92	3.62	136.38	14.65
		43.6%	56.4%	28.20%	48.72%	23.08%						

[a] *High Fear, High Attentional Control.* [b] *High Fear, Low Attentional Control.* [c] *Low Fear, High Attentional Control.* [d] *Low Fear, Low Attentional Control.*

and Anxiety as covariates. This analysis indicated no significant main effect, but a significant three-way interaction effect of Emotion valence by Fear level by Attentional Control level, $F_{(5,\ 151)} = 7.72$, $p < 0.01$, partial $\eta^2 = 0.05$ (Bonferroni correction applied). In order to understand the three-way interaction we completed two separate ANCOVAs, one for the angry bias scores and another for the happy bias scores.

The 2 × 2 ANCOVA (Age and Anxiety as covariates) for angry bias scores indicated a significant interaction effect of Fear and Attentional Control levels on bias scores, $F_{(3,\ 152)} = 5.58$, $p = 0.01$, partial $\eta^2 = 0.03$ (Bonferroni correction applied). No main effects of Fear, $F_{(3,\ 152)} = 1.22$, *ns*, Attentional Control, $F_{(3,\ 152)} = 0.05$, *ns*, Age, $F_{(3,\ 152)} = 0.002$, *ns*, or Anxiety, $F_{(3,\ 152)} = 0.20$, *ns*, reached significance[3]. As such, highly fearful children who also have high levels of attentional control seem to have weaker attentional biases toward threat, as compared to highly fearful children with low levels of attentional control (see **Table 2** for means and standard deviations).

Further we decomposed the interaction effect with follow-up *t*-tests. We looked at the main effect of fear on attentional biases toward angry faces as a function of attentional control. When comparing the low fear low attentional control group with the high fear, low attentional control group, for threat bias scores we observed a significant difference, $t_{(80)} = -3.03$, $p = 0.003$, $d = 0.68$ two-tailed. Inspecting the means from **Table 2**, we can see that children with high temperamental fear and low attentional control were significantly vigilant toward angry faces. When we looked at the other two groups and compared children with low fear and high attentional control to children with high fear and high attentional control, we observed a non-significant effect, $t_{(74)} = 0.59$, *ns*.

[3]The ANCOVA analysis was also conducted with the addition of Gender as a between variable (resulting in a 2 × 2 × 2 ANCOVA). However, results indicated no significant main effect of Gender, $F_{(5,\ 150)} = 0.40$, *ns*, or interaction effects of Gender with Fear and Attentional control, $F_{(5,\ 150)} = 0.70$, *ns*.

We also ran several one-sample *t*-tests in order to compare bias scores for each group to 0. When bias scores are significantly different from 0, they indicate a clear attentional bias. For the low fear, low attentional control group, the mean bias score was significantly different from 0, $t_{(38)} = -2.26$, $p = 0.02$, $d = 0.51$. The same was true for the high fear, low attentional control group $t_{(42)} = 2.06$, $p = 0.04$, $d = 0.45$. In the low fear high attentional control group, the mean bias score was not significantly different from 0, $t_{(52)} = 0.09$, *ns*. Also, the mean bias score did not significantly differ from 0 in the high fear, high attentional control group, $t_{(22)} = -0.58$, *ns*. Consequently, attentional biases appear to be present in the two groups of children that have low attentional control, at both high and low levels of fear. Specifically, children with high fear and low attentional control are significantly vigilant toward angry faces, whereas children with low fear and low attentional control present a significant attentional avoidance of angry faces. Children high in attentional control, with either low or high levels of fear, are not significantly biased in their attentional responses when confronted with an angry face (see **Figure 1**).

We conducted a second ANCOVA for the happy-neutral trials. We looked for possible effects of Fear and Attentional Control on bias scores for the happy-neutral stimuli, also controlling for the effects of Age and Anxiety. Results indicated no main effect of Fear, $F_{(3,\ 152)} = 0.004$, *ns*., Attentional Control, $F_{(3,\ 152)} = 2.99$, *ns*., and no interaction effect, $F_{(3,\ 152)} = 0.23$, *ns*. Also, the effects of Age $F_{(3,\ 152)} = 0.19$, *ns*., and Anxiety $F_{(3,\ 152)} = 0.24$, *ns*., did not reach significance. Therefore, it seems that the relation between fear, attentional control, and attentional biases is not a significant one in the case of happy faces.

Regression analyses

Because both fear and attentional control were measured on a continuous scale, we conducted an additional analysis based on hierarchical regression, in order to test the interaction between these two variables in predicting attentional biases toward angry faces. In addition, another potential difficulty in using ANCOVA

Table 2 | Mean threat reaction times for each condition and mean bias scores for the four groups (with standard deviations in parentheses).

Type of emotion	Reaction times		Attentional bias score
	Emotional congruent	Emotional incongruent	
HIGH FEAR, HIGH ATTENTIONAL CONTROL GROUP			
Angry	470.98 (69.92)	467.34 (63.97)	−3.64 (29.83)
Happy	463.02 (62.36)	467.76 (60.88)	4.74 (26.74)
HIGH FEAR, LOW ATTENTIONAL CONTROL GROUP			
Angry	489.30 (59.56)	497.83 (61.78)	8.53 (27.14)
Happy	491.91 (60.94)	490.43 (56.28)	−1.48 (24.52)
LOW FEAR, HIGH ATTENTIONAL CONTROL GROUP			
Angry	479.83 (72.47)	480.17 (77.82)	0.34 (25.24)
Happy	475.29 (73.24)	470.95 (74.75)	−4.34 (28.65)
LOW FEAR, LOW ATTENTIONAL CONTROL GROUP			
Angry	491.19 (83.29)	482.42 (85.14)	−8.77 (24.19)
Happy	486.62 (85.17)	489.37 (84.85)	2.75 (23.63)

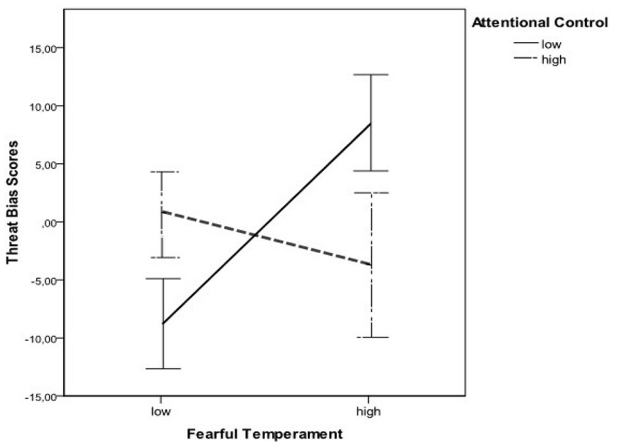


FIGURE 1 | Interaction effect of temperamental fear and temperamental attentional control on threat bias scores (bars represent values of standard errors).

arises from the use of correlated fear and attentional control measures ($r = -0.30$ in this sample), which may lead to inflated ANCOVA interaction if dichotomous groups are formed through median splits (Derryberry and Reed, 2002).

Therefore, hierarchical regression has the advantage of overcoming the problems of dichotomization of continuous variables based on median split procedures (Cohen et al., 2003). Following Aiken and West's (1991) guidelines, all variables were first centered and the interaction term (Fear × Attentional control) was computed as the multiplicative product of these two centered variables. Age and Anxiety were first entered. Fear was entered in the second step, followed by the Attentional control in the third step. The interaction term was entered in the fourth step.

Consistent with the results from ANCOVA, this analysis (see **Table 3**) yielded a significant Fear x Attentional control interaction on step forth ($b = -1.39$, $p = 0.01$, $f^2 = 0.06$). However, steps 1–4 were not significant (all $ps > 0.05$). We examined the particular form of this interaction by plotting the regression of threat bias scores on temperamental fear at high (one standard deviation above the mean), medium, and low (one standard deviation below the mean) levels of fear and attentional control. As shown in **Figure 2**, the slope was significantly different from zero only at low levels of attentional control, $t_{(154)} = 2.73$, $p < 0.01$. More specifically, there was a significant positive association between fear and attentional biases toward angry faces only for children with low attentional control. At high or medium values the slopes were not significantly different from zero, $t_{(154)} = -55$, $p = 0.57$ and $t_{(154)} = 1.63$, $p = 0.10$. These results indicate that there is no significant relation between temperamental fear and attentional vigilance toward threatening stimuli for children with good abilities for attentional control.

Consistent with the results from ANCOVA, no significant results were found for fear ($b = 2.81$, $p = 0.41$), attentional control ($b = -30$, $p = 0.52$), or interaction term (Fear × Attentional control $b = 0.70$, $p = 0.21$) in explaining attentional biases toward happy faces.

Differentiating engagement and difficulty to disengage in the Dot-probe task. As Koster and colleagues have pointed out, by comparing neutral-neutral trials in the Dot-probe task separately to congruent, respectively incongruent emotional-neutral trials, one could separate two components of attentional biases taking the form of heightened vigilance toward threat: faster engagement vs. difficulty of disengagement (Koster et al., 2004). Therefore, we also computed engagement and disengagement bias scores and conducted two separate hierarchical regressions to pinpoint the attentional mechanism responsible for the tendency of high fear low attentional control children to manifest greater attentional biases to threat. Engagement bias score reflects a faster response on congruent angry trials compared to neutral trials. This faster response is considered to show that individuals were preferentially holding their attention at the location of the angry face. Disengagement bias score reflects higher reaction times on the incongruent angry trials, due to the time needed to shift attention from the angry to the neutral location. We employed the following formulas in order to calculate these bias scores:

Engagement score = (Neutral-neutral trials reaction time) − (Congruent trials reaction time)
Disengagement score = (Incongruent trials reaction time) − (Neutral-neutral trials reaction time)

We also conducted two regression analyses for the engagement score, respectively disengagement score.

The regression analysis for the engagement score yielded a significant Fear × Attentional control interaction on the fourth step ($b = -1.37$, $p = 0.01$, $f^2 = 0.06$). Also, steps 1–4 were not significant (all $ps > 0.05$). In the regression analysis for the disengagement score no significant results were found for fear ($b = -3.14$, *ns*), attentional control ($b = -0.15$, *ns*), or interaction term (Fear × Attentional control $b = 0.001$, *ns*) in explaining difficulty to disengage from angry faces. Therefore, it seems that only the faster engagement to threat is implicated in the variations of attentional biases as a factor of temperamental fear and temperamental attentional control in the current study.

Table 3 | Summary of hierarchical regression analysis for variables predicting attentional biases toward angry faces.

Predictor	Δ R^2	*SE b*	*B*
Step 1	0.01		
Age (in months)		0.15	0.01
Anxiety		0.14	0.20
Step 2	0.02		
Fear		3.48	3.24
Step 3	0.02		
Attentional control		0.49	−0.39
Step 4	0.06		
Fear × attentional control		0.56	−1.39*
Total R^2	0.11		

*N = 158; *p < 0.05.*

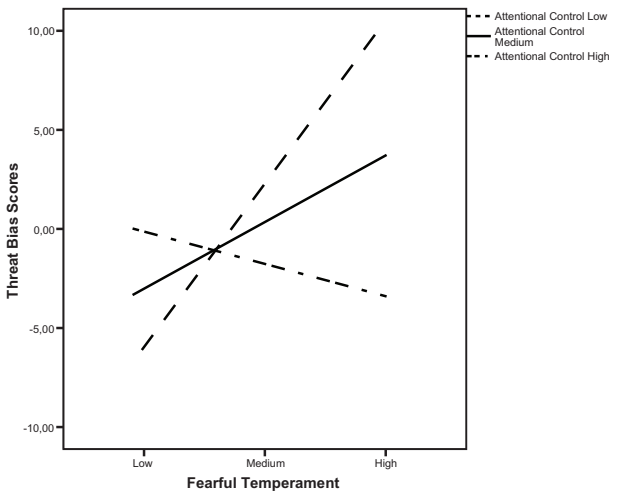


FIGURE 2 | The regression of threat bias scores on fearful temperament and attentional control (straight lines represent expected scores).

DISCUSSION

The present study aimed to investigate the effects of individual differences in temperamental fear and temperamental attentional control on attention allocation toward threat. Specifically, we analyzed the role of attentional control in regulating threat-related attentional biases in children with high levels of temperamental fear.

With regard to the main effects of both temperamental variables on attentional biases toward angry faces, neither fear nor attentional control was significantly related to attentional biases. However, consistent with our prediction, we found a significant interaction effect of fear and attentional control on attentional allocation toward threat. In particular, children with low levels of attentional control and high levels of fear displayed a stronger vigilance bias toward angry faces, compared to children who have low levels of attentional control and also low levels of fear. This vigilance seems to be underlained by an enhanced engagement of attention by angry faces, as it is proved by our additional regression analysis conducted on the two components of bias scores identified following Koster et al. (2004). This result is consistent with theoretical accounts on attentional biases toward threat, that generally link both the automatic/pre-attentional threat detection mechanism and the disruption of effortful strategies such as temperamental attentional control, with enhanced engagement of attention by angry faces (Beck and Clark, 1997; Mathews and MacKintosh, 1998; Mogg and Bradley, 1998; Cisler and Koster, 2010). But the lack of a significant difference between children with low fear, high attentional control and children with high fear, high attentional control indicates that when attentional control is increased, high levels of fear are not associated with biased attention toward angry faces. Also, the one-sample *t*-tests analysis comparing bias scores to 0 showed that children with high fear and low attentional control were indeed significantly vigilant toward angry faces. In addition, this analysis demonstrated that the group of children with low fear and low attentional control displayed a significant bias away from angry faces. Therefore, it seems that low attentional control is a key variable associated with biased attentional allocation in relation to angry facial expressions. We also noted that children high in both fear and attentional control did not show a significant bias.

The significant interaction between fear and attentional control replicates earlier findings, showing that attentional biases toward threat were present only in children who had both high levels of negative affectivity, such as fear, and low levels of regulative temperamental traits, such as attentional control (Helzer et al., 2009; Lonigan and Vasey, 2009). Our results revealed that, in highly fearful children the modulating role of high temperamental attentional control is reflected by a tendency to display attentional avoidance in the presence of threatening information. This attentional avoidance may involve a substantial voluntary component, relative to attentional vigilance toward threatening stimuli that accompanied the response of highly fearful children with low abilities of attentional control. Fearful children might be thought of as particularly vulnerable to automatically orient their attention toward threatening stimuli in the environment. But our results, in line with previous findings mentioned above, point out that in circumstances when attentional control can be employed to inhibit the orientation of attentional resources toward threat, only a subset of fearful children (those with low attentional control) go on to exhibit this reactive attentional response.

To our knowledge, this study is one of the first to investigate the possibility that individual differences in attentional control might modulate threat-related biases in fearful children, when ecological stimuli, such as emotional faces, are presented. A similar approach with a pictorial Dot-probe detection task, but with stimuli selected from the International Affective Picture System (Lang et al., 2005), is that of Vervoort et al. (2011). These authors examined the links between reactive temperament (negative affectivity as a composite factor), regulative temperament (effortful control as a composite factor), attentional biases and internalizing problems, in adolescents with and without anxiety disorders. Of direct relevance to our study, initial attentional biases (e.g., when stimulus duration was 500 ms) were predicted neither by the negative affectivity—effortful control interaction, nor by a main effects model, in either group. In addition, when stimulus duration was 1250 ms, higher levels of effortful control were related to attentional biases away from threat, but only in the non-anxious group, whereas in the anxious group effortful control had almost no influence on attentional biases. Our results complement these data by demonstrating that, in a non-clinical sample of children, the regulative temperamental trait, here more specifically assessed as attentional control influenced the threat-related attentional biases pattern, as children with both high levels of attentional control and also high levels of fear manifested a pattern of attentional avoidance in relation to threatening stimuli. The added value of the present results is reflected in the finding that attentional control influences initial attentional biases, at least toward angry faces, since a stimulus duration of 500 ms is assumed to reflect early initial attention (Bradley et al., 2000). We observed attentional bias scores significantly different from 0 in the two groups of children characterized by low levels of attentional control. Thus, our data support the conclusion that individual differences in attentional control have to be considered when investigating threat-related attentional biases in children with non-clinical anxiety. Not taking this variable into consideration might explain the divergent pattern of results obtained in previous studies (Lonigan and Vasey, 2009).

Another important aspect was the lack of any moderating effects of age or anxiety level. The lack of an age effect is similar to results of other studies with different age groups (e.g., Hadwin et al., 2009; Lonigan and Vasey, 2009). However, it is divergent from the results of a recent study on trait anxiety in children, showing a main effect of age on emotional processing and a moderating effect of age on attentional biases for negative stimuli, in a modified Stroop task (Reinholdt-Dunne et al., 2012). Interestingly, their study included a wider age range (7–14) than ours (9–14), therefore the lack of age-related effects in our data does not rule out the possibility of differential emotional or, more specifically, threat processing in younger children.

In the present study, anxiety symptoms did not influence attentional bias scores. This is somewhat similar to the lack of attentional biases for threat in anxious children, reported by Reinholdt-Dunne and collaborators in the case of older children (mean age 11) (Reinholdt-Dunne et al., 2012). One possibility

is that the lack of association between anxiety and attentional biases was due to the non-clinical sample involved in the current study, a point also made by Reinholdt-Dunne and colleagues in reference to their results. This explanation is supported by the failure of some previous studies conducted with non-clinical samples to find evidence for an association between high levels of anxiety (e.g., high levels of trait anxiety) and biases toward threat (Eschenbeck et al., 2004; Helzer et al., 2009). Also, there are studies which suggest that moderate to severe levels of clinical anxiety in children are reliably associated with increased attentional biases toward angry faces relative to neutral faces (Waters et al., 2010a). An alternative explanation is that, for children with non-clinical anxiety, the emotional reactivity related to anticipation of stress, derived from temperament fear, might influence the direction of attention to threatening information more than anxiety. This finding requires replication by including the assessment of both reactive temperamental fear and anxiety symptoms in future studies that investigate attentional biases with non-clinical samples. Moreover, it should be mentioned that our study was designed to evaluate whether there would be group differences in attentional biases, as a function of temperamental traits and their interaction. Thus, we did not preselect our sample based on extreme anxiety scores. Therefore, the absence of a relation between anxiety and bias scores from the present study does not indicate that the full model proposed by Lonigan and colleagues regarding the relations between temperament, attentional biases and anxiety is not plausible. Future studies should analyze the stability and change over time of these relations in a longitudinal design. However, the present findings provide evidence only for a relation between temperament and attentional biases.

In our study we also examined attentional biases for happy faces. However, we did not formulate any specific predictions regarding the direction of attentional processes for happy faces, given that some studies conducted with children (Waters et al., 2008) have found a bias toward this kind of stimuli, whereas others have not (Telzer et al., 2008). The analysis of happy-neutral trials revealed no relation between attentional biases for happy faces and temperamental traits. This result is in line with previous studies that revealed no attentional biases in relation to happy facial expressions in anxious youths or in children with underlying anxiety predispositions (Roy et al., 2008; Telzer et al., 2008).

There are several limitations to be considered when interpreting our current findings. First, temperamental traits and attentional biases were assessed concurrently. Therefore, no conclusion can be inferred regarding the directionality of the observed effects. From a developmental perspective, it is important to shed light on the specific ways these variables influence each other, so that longitudinal studies assessing these factors will be needed. Second, our study investigated only one part of the model formulated by Lonigan and his collaborators. In order to adequately test the full model, data should be collected longitudinally. For example, future studies should analyze the impact of attentional biases on anxiety symptoms in children with certain temperamental characteristics. Third, as this study included only children without anxiety disorders, the observed effects cannot be generalized to clinically anxious children, for whom the nature of attentional processes and their relations with temperamental traits may be different (Vervoort et al., 2011). Moreover, given that we used self-report instruments for both temperamental fear and attentional control, it would be important to complement such measurements in future studies, for example with a behavioral task for attentional control. In the present study, we tried to overcome the problem of correlated fear and attentional control measures by also conducting a hierarchical regression in order to analyze our data. An additional aspect to be noted here is connected to the methodological weaknesses of the Dot-probe paradigm. It has been pointed out that reaction time effects in this task could be due to behavioral interference rather than to attentional phenomenon *per se*, especially at longer stimulus durations (e.g., Wolters et al., 2012). We tried to control for such confounds by running a preliminary analysis, to compare reaction times on all neutral faces trials to reaction times on all angry, respectively all happy faces trials, which showed no significant differences. However, it remains open to discussion whether the generally accepted calculation of bias scores in this task can accurately differentiate between attentional vigilance and avoidance (with positive bias scores indicating vigilance and negative ones indicating avoidance). This is because, during the 500 ms stimulus presentation interval, several shifts of attention are possible (Weierich et al., 2008). Thus, without systematic variation in display time and/or eye movements monitoring, it is virtually impossible to be certain what reaction times stand for, in terms of attentional vigilance vs. avoidance, at the end of the 500 ms interval. Therefore, conclusions regarding the presence of attentional biases of vigilance toward threat as opposed to avoidance of threat are to be regarded with caution. It is very important that future reaction times studies strive to provide better control in pinpointing the time course of attentional shifts.

In conclusion, despite the inherent limitations, the present results point to the importance of studying threat-related attentional biases in relation to temperamental traits. Results indicate that heightened vigilance toward angry faces is characteristic only of children with high fear and low attentional control.

The present study indicates that temperamentally-based attentional control plays a regulative role, modulating reactivity that characterizes temperamental fearfulness. Therefore, based on our data, we advance the hypothesis that attentional control can be seen as a possible early protective factor for the development of attentional biases toward threat, and further for the manifestation of anxiety problems. Future work, using a longitudinal design with both clinical and non-clinical samples, is required to examine this hypothesis.

ACKNOWLEDGMENTS

This work was supported by a grant of the Ministry of National Education, CNCS—UEFISCDI, project number PN-II-ID-PCE-2012-4-0668 awarded to the second author.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00922/abstract

REFERENCES

Aiken, L. S., and West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th Edn*. Washington, DC: American Psychiatric Press.

Anderson, M., Nettelbeck, T., and Barlow, J. (1997). Reaction time measures of speed of processing: Speed of response selection increases with age but speed of stimulus categorization does not. *Br. J. Dev. Psychol.* 15, 145–157. doi: 10.1111/j.2044-835X.1997.tb00731.x

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., and van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol. Bull.* 133, 1–24. doi: 10.1037/0033-2909.133.1.1

Beck, A. T., and Clark, D. A. (1997). An information processing model of anxiety: automatic and strategic processes. *Behav. Res. Ther.* 35, 49–58. doi: 10.1016/S0005-7967(96)00069-1

Bell, M. A., and Calkins, S. D. (2012). "Attentional control and emotion regulation in early development," in *Cognitive Neuroscience of Attention, 2nd Edn*., ed M. I. Posner (New York, NY: Guilford Press), 322–330.

Benga, O., Tincas, I., Visu-Petra, L., Pitica, I., and Susa, G. (in press). Investigating the structure of anxiety symptoms among Romanian school aged children using the Spence Child Anxiety Scale. *Cogn. Brain Behav.*

Bradley, B. P., Mogg, K., Falla, S. L., and Hamilton, L. R. (1998). Attentional bias for threatening facial expressions in anxiety: Manipulation of stimulus duration. *Cogn. Emot.* 12, 737–753. doi: 10.1080/026999398379411

Bradley, B. P., Mogg, K., and Millar, N. H. (2000). Covert and overt orienting of attention to emotional faces in anxiety. *Cogn. Emot.* 14, 789–808. doi: 10.1080/026999300501566636

Brotman, M. A., Rich, B. A., Schmajunk, M., Reising, M., Monk, C. S., Dickstein, D. P., et al. (2007). Attentional biases to threat faces in children with bipolar disorder and comorbid lifetime anxiety disorders. *Biol. Psychiatry* 61, 819–821. doi: 10.1016/j.biopsych.2006.08.021

Calkins, S. D., and Fox, N. A. (2002). Self-regulatory processes in early personality development: a multilevel approach to the study of childhood social withdrawal and aggression. *Dev. Psychopathol.* 14, 477–498. doi: 10.1017/S095457940200305X

Cisler, J. M., and Koster, E. H. (2010). Mechanisms of attentional biases toward threat in anxiety disorders: an integrative review. *Clin. Psychol. Rev.* 30, 203–216. doi: 10.1016/j.cpr.2009.11.003

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Science, 3rd Edn*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dalgleish, T., Taghavi, R., Neshat-Doost, H., Moradi, A., Canterbury, R., and Yule, W. (2003). Patterns of processing bias for emotional information across clinical disorders: a comparison of attention, memory and prospective cognition in children and adolescents with depression, generalized anxiety and posttraumatic stress disorder. *Clin. Child Adolesc. Psychol.* 32, 10–21. doi: 10.1207/S15374424JCCP3201_02

Derryberry, D., and Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *J. Abnorm. Psychol.* 111, 225–236. doi: 10.1037/0021-843X.111.2.225

Eisenberg, N., Guthrie, I. K., Fabes, R. A., Shepard, S., Losoya, S., Murphy, B., et al. (2000). Prediction of elementary school children's externalizing problem behaviors from attentional and behavioral regulation and negative emotionality. *Child Dev.* 71, 1367–1382. doi: 10.1111/1467-8624.00233

Ekman, P., and Friesen, W. (1976). *Pictures of Facial Affect*. Palo Alto, CA: Consulting Psychology Press.

Ellis, L. K., and Rothbart, M. K. (2001). "Revision of the early adolescent temperament questionnaire," in *Poster presented at the 2001 Biennial Meeting of the Society for Research in Child Development* (Minneapolis, MN).

Eschenbeck, H., Kohlmann, C. W., Heim-Dreger, U., Koller, D., and Leser, M. (2004). Processing bias and anxiety in primary school children: a modified emotional Stroop colour-naming task using pictorial facial expressions. *Psychol. Sci.* 46, 451–465.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., and Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* 14, 340–347. doi: 10.1162/089892902317361886

Fox, N. A., Henderson, H. A., Marshall, P. J., Nichols, K. E., and Ghera, M. M. (2005).Behavioral inhibition: linking biology and behavior within a developmental framework. *Annu. Rev. Psychol.* 56, 235–262. doi: 10.1146/annurev.psych.55.090902.141532

Hadwin, J. A., Donnelly, N., Richards, A., French, C. C., and Patel, U. (2009). Childhood anxiety and attention to emotion faces in a modified Stroop task. *Br. J. Dev. Psychol.* 27, 487–494. doi: 10.1348/026151008X315503

Hardee, J. E., Benson, B. E., Bar-Haim, Y., Mogg, K., Bradley, B. P., Chen, G., et al. (2013). Patterns of neural connectivity during an attention bias task moderate associations between early childhood temperament and internalizing symptoms in young adulthood. *Biol. Psychiatry* 74, 273–279. doi: 10.1016/j.biopsych.2013.01.036

Helzer, E. G., Connor-Smith, J. K., and Reed, M. A. (2009). Traits, states, and attentional gates: temperament and threat relevance as predictors of attentional bias to social threat. *Anxiety Stress Coping* 22, 57–76. doi: 10.1080/10615800802272244

Henderson, H. A., and Wachs, T. D. (2007). Temperament theory and the study of cognition-emotion interactions across development. *Dev. Rev.* 27, 396–427. doi: 10.1016/j.dr.2007.06.004

Iida, Y., Miyazaki, M., and Uchida, S. (2010). Developmental changes in cognitive reaction time of children aged 6–12 years. *Eur. J. Sport Sci.* 10, 151–158. doi: 10.1080/17461390903515162

In-Albon, T., Kossovsky, J., and Schneider, S. (2010). Vigilance and avoidance of threat in the eye movements of children with separation anxiety disorder. *J. Abnorm. Child Psychol.* 38, 225–235. doi: 10.1007/s10802-009-9359-4

Ingram, R. E., and Price, J. M. (2010). *Vulnerability to Psychopathology: Risk Across the Lifespan*. New York, NY: The Guilford Press

Kagan, J., Reznick, J. S., and Snidman, N. (1987). The physiology and psychology of behavioral inhibition in children. *Child Dev.* 58, 1459–1473.

Koster, E. H. W., Crombez, G., Verschuere, B., and De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: differentiating vigilance and difficulty to disengage. *Behav. Res. Ther.* 42, 1183–1192. doi: 10.1016/j.brat.2003.08.001

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2005). *International Affective Pictures System (IAPS): Affective Ratings of Pictures and Instruction Manual. Technical Report A-6*. Gainesville, FL: University of Florida.

Lonigan, C. J., and Vasey, M. W. (2009). Negative affectivity, effortful control, and attention to threat-relevant stimuli. *J. Abnorm. Child Psychol.* 37, 387–399. doi: 10.1007/s10802-008-9284-y.

Lonigan, C. J., Vasey, M. W., Phillips, B. M., and Hazen, R. A. (2004). Temperament, anxiety, and the processing of threat-relevant stimuli. *J. Clin. Child Adolesc. Psychol.* 33, 8–20. doi: 10.1207/S15374424JCCP3301_2

Martin, M., and Jones, G. V. (1995). Integral bias in the cognitive processing of emotionally linked pictures. *Br. J. Psychol.* 86, 419–435. doi: 10.1111/j.2044-8295.1995.tb02761.x

Mathews, A., and MacKintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cognit. Ther. Res.* 22, 539–550. doi: 10.1023/A:1018738019346

Mogg, K., and Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behav. Res. Ther.* 36, 809–848. doi: 10.1016/S0005-7967(98)00063-1

Mogg, K., and Bradley, B. P. (1999). Some methodological issues in assessing attentional biases for threatening faces in anxiety: a replication study using a modified version of the probe detection task. *Behav. Res. Ther.* 37, 595–604. doi: 10.1016/S0005-7967(98)00158-2

Mogg, K., Garner, M., and Bradley, B. P. (2007). Anxiety and orienting of gaze to angry and fearful faces. *Biol. Psychol.* 76, 163–169. doi: 10.1016/j.biopsycho.2007.07.005

Monk, C. S., Nelson, E. E., McClure, E. B., Mogg, K., Bradley, B. P., Leibenluft, E., et al. (2006). Ventrolateral prefrontal cortex activation and attentional bias in response to angry faces in adolescents with generalized anxiety disorder. *Am. J. Psychiatry* 163, 1091–1097. doi: 10.1176/appi.ajp.163.6.1091

Muris, P., De Jong, P. J., and Engelen, S. (2004). Relationship between neuroticism, attentional control, and anxiety disorders symptoms in non-clinical children. *Pers. Indiv. Differ.* 37, 789–797. doi: 10.1016/j.paid.2003.10.007

Muris, P., and Meesters, C. (2009). Reactive and regulative temperament in youth: psychometric evaluation of the Early Adolescent Temperament Questionnaire-Revised. *J. Psychopathol. Behav. Assess.* 31, 7–19. doi: 10.1007/s10862-008-9089-x

Muris, P., Meesters, C., and Rompelberg, L. (2007). Attention control in middle childhood: relations to psychopathological symptoms and threat perception distorsions. *Behav. Res. Ther.* 45, 997–1010. doi: 10.1016/j.brat.2006.07.010

Oehlberg, K. A., Revelle, W., and Mineka, S. (2012). Time-course of attention to negative stimuli: negative affectivity, anxiety, or dysphoria? *Emotion* 12, 943–959. doi: 10.1037/a0027227

Pine, D. S., Helfinstein, S. M., Bar-Haim, Y., Nelson, E., and Fox, N. A. (2009). Challenges in developing novel treatments for childhood disorders: lessons from research on anxiety. *Neuropsychopharmacol. Rev.* 34, 213–228. doi: 10.1038/npp.2008.113

Posner, M. I., and DiGirolamo, G. J. (1998). "Executive attention: conflict, target detection, and cognitive control," in *The Attentive Brain,* ed R. Parasuraman (Cambridge, MA: The MIT Press), 401–423.

Posner, M. I., and Rothbart, M. K. (2000). Developing mechanisms of self regulation. *Dev. Psychopathol.* 12, 427–441. doi: 10.1017/S0954579400003096

Posner, M. I., and Rothbart, M. K. (2007). *Educating the Human Brain*. Washington, DC: American Psychological Association.

Reinholdt-Dunne, M. L., Mogg, K., Esbjorn, B. H., and Bradley, B. P. (2012). Effects of age and anxiety on processing threat cues in healthy children. *J. Exp. Psychopathol.* 3, 30–41. doi: 10.5127/jep.019611

Richards, A., Richards, L. C., and McGeeney, M. (2000). Anxiety-related Stroop interference in adolescents. *J. Gen. Psychol.* 127, 327–333. doi: 10.1080/00221300009598587

Rothbart, M., and Bates, J. (2006). "Temperament," in *Handbook of Child Psychology: Social, Emotional, and Personality Development, 6th Edn.,* Vol. 3, eds N. Eisenberg, W. Damon, and L. M. Richard (Hoboken, NJ: John Wiley & Sons Inc), 99–166.

Rothbart, M. K., and Derryberry, D. (1981). "Development of individual differences in temperament," in *Advances in Developmental Psychology,* Vol. 1, eds M. E. Lamb and A. L. Brown (Hillsdale, NJ: Lawrence Erlbaum Associates), 37–86.

Roy, A. K., Vasa, R. A., Bruck, M., Mogg, K., Bradley, B. P., Sweeney, M., et al. (2008). Attention bias toward threat in pediatric anxiety disorders. *J. Am. Acad. Child Adolesc. Psychiatry* 47, 1189–1196. doi: 10.1097/CHI.0b013e3181825ace

Rueda, M. R., Posner, M. R., and Rothbart, M. K. (2005). The development of executive attention: contributions to the emergence of self-regulation. *Dev. Neuropsychol.* 28, 573–594. doi: 10.1207/s15326942dn2802_2

Schmidt, L. A., Fox, N. A., Rubin, K. H., Sternberg, E. M., Gold, P. W., Smith, C. C., et al. (1997). Behavioral and neuroendocrine responses in shy children. *Dev. Psychobiol.* 30, 127–140. doi: 10.1002/(SICI)1098-2302(199703)30:2<127::AID-DEV4>3.0.CO;2-S

Simonds, J., Kieras, J. E., Rueda, M. R., and Rothbart, M. K. (2007). Effortful control, executive attention, and emotional regulation in 7-10-year-old children. *Cogn. Dev.* 22, 474–488. doi: 10.1016/j.cogdev.2007.08.009

Spence, H. S. (1998). A measure of anxiety symptoms among children. *Behav. Res. Ther.* 36, 545–566. doi: 10.1016/S0005-7967(98)00034-5

Spence, H. S., Barrett, P. M., and Turner, C. M. (2003). Psychometric properties of the Spence Children's Anxiety Scale with young adolescents. *J. Anxiety Disord.* 17, 605–625. doi: 10.1016/S0887-6185(02)00236-0

Susa, G., Pitica, I., Benga, O., and Miclea, M. (2012). The self regulatory effect of attentional control in modulating the relationship between attentional biases toward threat and anxiety symptoms in children. *Cogn. Emot.,* 26, 1069–1083. doi: 10.1080/02699931.2011.638910

Szpunar, M., and Young, A. R. (2012). Information processing biases in behaviorally inhibited children: response to threat and novelty. *Child Health Educ.* 4, 47–63.

Telzer, E. H., Mogg, K., Bradley, B. P., Mai, X., Ernst, M., Pine, D. S., et al. (2008). Relationship between trait anxiety, prefrontal cortex and attention bias to angry faces in children and adolescents. *Biol. Psychol.* 79, 216–222. doi: 10.1016/j.biopsycho.2008.05.004

Tottenham, N., Hare, T. A., and Casey, B. J. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Front. Psychol.* 2:39. doi: 10.3389/fpsyg.2011.00039

Tottenham, N., Tanaka, J., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgements from untrained research participants. *Psychiatry Res.* 168, 242–249. doi: 10.1016/j.psychres.2008.05.006

Vasey, M. V., Daleiden, E. L., Williams, L. L., and Brown, L. M. (1995). Biased attention in childhood anxiety disorders: a preliminary study. *J. Abnorm. Child Psychol.* 23, 267–279. doi: 10.1007/BF01447092

Vasey, M. W., El-Hag, N., and Daleiden, E. L. (1996). Anxiety and the processing of emotionally threatening stimuli: distinctive patterns of selective attention among high- and low- test- anxious children. *Child Dev.* 67, 1173–1185. doi: 10.1111/j.1467-8624.1996.tb01789.x

Vervoort, L. V., Wolters, L. H., Hogendoorn, S. M., Prins, P. J., de Haan, E., and Hartman, C. A. (2011). Temperament, attentional processes, and anxiety: diverging links between adolescents with and without anxiety disorders? *J. Clin. Child Adolesc. Psychol.* 40, 144–155. doi: 10.1080/15374416.2011.533412

Waters, A. M., Henry, J., Mogg, K., Bradley, B. P., and Pine, D. S. (2010a). Attentional bias toward angry faces in childhood anxiety disorders. *J. Behav. Ther. Exp. Psychiatry* 41, 158-164. doi: 10.1016/j.jbtep.2009.12.001

Waters, A. M., Kokkoris, L. L., Mogg, K., Bradley, B. P., and Pine, D. S. (2010b). The time course of attentional bias for emotional faces in anxious children. *Cogn. Emot.* 24, 1173–1181. doi: 10.1080/02699930903274355

Waters, A. M., Mogg, K., Bradley, B. P., and Pine, D. S. (2008). Attentional bias for emotional faces in children with generalized anxiety disorder. *J. Am. Acad. Child Adolesc. Psychiatry* 47, 435–442. doi: 10.1097/CHI.0b013e3181642992

Weierich, M. R., Treat, T. A., and Hollingworth, A. (2008). Theories and measurement of visual attentional processing in anxiety. *Cogn. Emot*. 22, 985–1018. doi: 10.1080/02699930701597601

Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., Wright, C. I., and Rauch, S. L. (2001). A functional MRI study of human amygdala responses to facial expressions of fear versus anger. *Emotion* 1, 70–83. doi: 10.1037/1528-3542.1.1.70

White, L. K., Helfinstein, S. M., and Fox, N. A. (2010). "Temperamental factors associated with the acquisition of information processing biases and anxiety," in *Information Processing Biases in Child and Adolescent Anxiety,* eds J. Hadwin and A. Field (Oxford: Wiley-Blackwell), 233–252.

Williams, J. M., Watts, F. N., MacLeod, C., and Mathews, A. (1988). *Cognitive Psychology and Emotional Disorders.* Chichester, England: Wiley.

Wolters, L. H., de Haan, E., Vervoort, L., Hogendoorn, S. M., Boer, F., and Prins, P. J. M. (2012). The time-course of threat processing in children: a temporal dissociation between selective attention and behavioral interference. *Anxiety Stress Coping* 25, 259–273. doi: 10.1080/10615806.2011.581278

Yzerbyt, V. Y., Muller, D., and Judd, C. M. (2004). Adjusting researchers' approach to adjustment: on the use of covariates when testing interactions. *J. Exp. Soc. Psychol.* 40, 424–431. doi: 10.1016/j.jesp.2003.10.001

Zhou, Q., Chen, S. H., and Main, A. (2012). Commonalities and differences in the study of children's effortful control and executive function: a call for an integrated model of self-regulation. *Child Dev. Perspect.* 6, 112–121. doi: 10.1111/j.1750-8606.2011.00176.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2014; accepted: 02 August 2014; published online: 25 August 2014.

*Citation: Susa G, Benga O, Pitica I and Miclea M (2014) Child temperamental reactivity and self-regulation effects on attentional biases. Front. Psychol. **5**:922. doi: 10.3389/fpsyg.2014.00922*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Susa, Benga, Pitica and Miclea. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 19 May 2014
doi: 10.3389/fpsyg.2014.00447

Executive function and food approach behavior in middle childhood

Karoline Groppe* and Birgit Elsner

Department of Psychology, Developmental Psychology, University of Potsdam, Potsdam, Germany

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Masatoshi Katagiri, University of Toyama, Japan
Yoshifumi Ikeda, Tokyo Gakugei University, Japan

****Correspondence:***
Karoline Groppe, Department of Psychology, Developmental Psychology, University of Potsdam, Karl-Liebknecht-Street 24/25, 14476 Potsdam, Germany
e-mail: karoline.groppe@uni-potsdam.de

Executive function (EF) has long been considered to be a unitary, domain-general cognitive ability. However, recent research suggests differentiating "hot" affective and "cool" cognitive aspects of EF. Yet, findings regarding this two-factor construct are still inconsistent. In particular, the development of this factor structure remains unclear and data on school-aged children is lacking. Furthermore, studies linking EF and overweight or obesity suggest that EF contributes to the regulation of eating behavior. So far, however, the links between EF and eating behavior have rarely been investigated in children and non-clinical populations. First, we examined whether EF can be divided into hot and cool factors or whether they actually correspond to a unitary construct in middle childhood. Second, we examined how hot and cool EF are associated with different eating styles that put children at risk of becoming overweight during development. Hot and cool EF were assessed experimentally in a non-clinical population of 1657 elementary-school children (aged 6–11 years). The "food approach" behavior was rated mainly via parent questionnaires. Findings indicate that hot EF is distinguishable from cool EF. However, only cool EF seems to represent a coherent functional entity, whereas hot EF does not seem to be a homogenous construct. This was true for a younger and an older subgroup of children. Furthermore, different EF components were correlated with eating styles, such as responsiveness to food, desire to drink, and restrained eating in girls but not in boys. This shows that lower levels of EF are not only seen in clinical populations of obese patients but are already associated with food approach styles in a normal population of elementary school-aged girls. Although the direction of effect still has to be clarified, results point to the possibility that EF constitutes a risk factor for eating styles contributing to the development of overweight in the long-term.

Keywords: hot and cool executive function, eating behavior, food approach, overweight, middle childhood

INTRODUCTION

Self-regulation, which is one of the major achievements in early childhood, is facilitated through a variety of processes which are referred to as executive functions. Executive function (EF) has been found to be strongly (but not exclusively) linked to the prefrontal cortex (PFC; for a meta-analysis see Alvarez and Emory, 2006) and enables the control of thoughts, actions, and emotions (e.g., Zelazo et al., 2008) via a number of related but distinct subfunctions, including shifting, updating, and inhibition (Miyake et al., 2000). EF has long been considered to be a unitary, domain-general cognitive function with its subfunctions working together in a consistent fashion across different situations and content domains (e.g., Zelazo et al., 1997). However, this assumption was partly based on traditional theories emphasizing exclusively one facet of EF measured by relatively abstract, decontextualized problems. More recent research indicates that a different facet of EF is needed when a task involves the regulation of affect and/or motivation (Happaney et al., 2004; Hongwanishkul et al., 2005). Hence, a distinction has been proposed between cognitive "cool" EF, which is activated when solving abstract novel problems, and affective "hot" EF, which is required for problems demanding high affective involvement or flexible appraisals of the affective significance of a stimulus (Zelazo and Müller, 2002).

Evidence for the distinction of hot and cool EF in adults comes from lesion and neuro-imaging studies on diverging functions of different parts of the prefrontal cortex (PFC; Zelazo and Müller, 2002; Happaney et al., 2004). Whereas dorsolateral regions of the PFC (DL-PFC) are associated with cool demands, ventral or medial regions of the PFC (VM–PFC), which are strongly connected to the limbic system, are required for hot regulatory tasks. Furthermore, the distinction is supported by findings that impairments in hot EF can occur in the absence of impairments in cool EF, and vice versa (e.g., Bechara, 2004; Eslinger et al., 2004).

However, to date, empirical findings on hot and cool EF in children remain inconsistent, and further research on its development is needed (Zelazo and Carlson, 2012). There is some indication that changes in cool EF occur earlier than changes in hot EF (e.g., Prencipe et al., 2011), and some studies on preschool-aged children have found that hot and cool EF performance can be described by separate but correlated factors that show different developmental correlates, like academic achievement (e.g., Brock et al., 2009; Willoughby et al., 2011), symptoms

of ADHD and behavioral problems, as well as social competence (Sonuga-Barke et al., 2003; Dalen et al., 2004; Smith-Donald et al., 2007). Other studies, however, have found important differences within hot EF tasks, challenging the assumption of a homogeneous hot factor (e.g., Hongwanishkul et al., 2005; Prencipe et al., 2011). Yet other studies found that hot and cool EF do not reflect different factors, but rather belong to a unitary construct in childhood (e.g., Allan and Lonigan, 2011; Wiebe et al., 2011). Some of this inconsistency may come from methodological problems, for instance, most of these studies did not account for the assumption that hot and cool EF are distinct but correlated processes in using either principal-component analyses or varimax rotations in their factor analyses (Willoughby et al., 2011). Moreover, research has to date focused on children younger than 7 years of age, and it might be that the distinction between hot and cool EF emerges later in the course of development, with an increasing functional specialization of neural systems (Johnson, 2011; Zelazo and Carlson, 2012).

To shed further light on the development of EF, the first aim of the present study was to examine whether EF measures can be divided into a hot and a cool factor or whether they correspond to a unitary construct in middle childhood. Because of some evidence for a two-factor structure in younger children (e.g., Brock et al., 2009; Willoughby et al., 2011) and because of the ongoing functional specialization of the neural systems (Johnson, 2011; Zelazo and Carlson, 2012), we expected to find two separate but correlated factors for hot and cool EF. In addition, we tested the factor structure in younger vs. older children of our sample in order to detect age-related differences that may inform about EF development between ages 6 and 11. In particular, we hypothesized that the hot cool distinction might become more evident in older children.

The construct of EF has also received much attention in research on eating disorders and obesity. Overweight and obesity, as well as eating disorders like bulimia and binge-eating disorder, typically involve a dysregulation of eating behavior that points to a prefrontal dysfunction, such as impulsive eating patterns (Spinella and Lyke, 2004). Neurological research supports this interrelation in providing a link between PFC functioning and the control of eating behavior. Imaging studies suggest that the PFC, particularly the VM–PFC, plays a role in different aspects of eating, like affecting the reinforcing value of food, disinhibited eating, hunger, food choice, or weight maintenance (e.g., Tataranni et al., 1999; Appelhans, 2009; Volkow et al., 2009; Cohen et al., 2011; Maayan et al., 2011).

Especially one facet of EF has received further attention in the context of eating, namely the inhibition of dominant responses. Increased impulsivity and reduced inhibitory control are associated with less healthy food choice (e.g., Bryant et al., 2008; Jasinska et al., 2012), eating in response to negative emotional states or external food cues (e.g., Bekker et al., 2004; Elfhag and Morey, 2008) as well as with binge eating (see Fischer et al., 2008; Waxman, 2009 for reviews) and a higher BMI (e.g., Nederkoorn et al., 2006; Batterink et al., 2010).

Furthermore, impairments in various aspects of hot and cool EF have been reported for overweight or obese individuals as compared to normal weight controls, independent of associated medical conditions (see Smith et al., 2011 for a review). For obese children and adolescents (4–18 years), 8 in 9 studies indicate deficits in set shifting, inhibition, working memory, attention, or affective decision-making (Smith et al., 2011). Additionally, there is a link between ADHD and being overweight indicating that EF deficits, as a symptom of ADHD, might favor overeating behaviors (see Cortese et al., 2008; Dempsey et al., 2011, for reviews).

To sum up, results from different research disciplines strongly suggest an association between EF and eating behavior. However, the topic has mostly been examined from a clinical perspective of eating disorders or obesity with the focus on EF deficits in overweight populations compared to controls. The few studies covering EF in relation to eating behavior (and not solely BMI) were limited to examining only inhibition in again mostly clinical populations of adults (e.g., Elfhag and Morey, 2008; Waxman, 2009). To our knowledge, so far, only one study has investigated associations between a broad range of EF and different eating styles in a population sample of adults, and this study reported associations between increased dysexecutive traits and disinhibited eating or greater food cravings (Spinella and Lyke, 2004). This points to a link between EF and eating, even in normal populations, suggesting that eating disorders or obesity represent only the extremes of a normal continuum of eating behavior. Moreover, except for the few studies on obese children (Smith et al., 2011), research on EF and eating or weight issues has almost exclusively focused on adult or adolescent populations. Yet, already children show variation in the extent to which they show food approach behavior, such as food responsiveness, emotional overeating, enjoyment of food, desire to drink, or external eating (Wardle et al., 2001; Sleddens et al., 2008). Illuminating early correlates of such eating behavior that put children at risk for higher weight gain would be of great importance for the prevention of overweight, especially considering the growing prevalence and serious consequences of being overweight and obese (Ogden et al., 2006; Moß et al., 2007).

Therefore, the second aim of the present study was to examine how hot and cool EF are associated with different eating styles that put children at risk of becoming overweight. We expected to find negative associations, i.e., difficulties in self-control, seen in lower levels of hot and cool EF, should co-occur with a higher level of various food approach behaviors in our sample of children in middle childhood. Because other studies have found gender effects for correlates of body weight, such as personality factors (e.g., Brummett et al., 2006; Armon et al., 2013) possible moderations by gender were tested exploratively.

METHODS

PARTICIPANTS

A total of 1657 children (52.1% girls) aged 6–11 years ($M = 8.3$ years, $SD = 0.95$, $Md = 8.4$ years) and their parents ($N = 1339$) participated in the study. Participants were recruited from 33 elementary schools from the federal state of Brandenburg (German school classes 1–3). Schools were preselected in terms of a representative variety of social backgrounds, as well as urban and rural areas.

Using the criteria of Kromeyer-Hauschild et al. (2001), 81.1% of the children were in the normal BMI range, 6.0% were underweight, 7.7% overweight, and 5.2% obese. This is broadly in line with other prevalence estimates. However, underweight as well as overweight children seemed to be slightly underrepresented (Kurth and Schaffrath Rosario, 2007).

MATERIAL

EF measures

Cool executive functions. The attention shifting component of cool EF was measured by the Cognitive Flexibility Task (Roebers et al., 2011; adapted from Zimmermann et al., 2002). Children were told to consecutively feed a plain and a colored fish that appeared simultaneously on the left and right side of a computer screen with randomly changing sides per trial. In order to feed the fish, the child needed to press one of two corresponding keys of a QWERTZ keyboard (the X-key for the left-side fish, the M-key for the right-side fish), remembering which kind of fish had been fed in the previous trial. There were 46 trials (22 switch-trials) separated by a short break that included positive feedback. The interstimuli intervals varied between 300 and 700 ms. The dependent variable for this task was the number of correct responses in the switch-trials (i.e., when the required answer set changed from a right-left to a right-right/left-left reaction, respectively).

The updating component of cool EF (monitoring and updating of working memory representations; Miyake et al., 2000) was assessed by the Digit Span Backwards Task (Petermann and Petermann, 2007). The child heard a sequence of numbers and had to verbally repeat it in reverse order. Each trial consisted of 2 sequences with the same number of digits. The experimenter started with a 2-digit-sequence and passed on to the next trial (one additional digit—except for trial 1 and 2 both consisting of only 2 digits) if at least one of the sequences in a trial had been answered correctly. The dependent variable was the total number of sequences correctly recalled.

The inhibition component of cool EF was measured by the Fruit Stroop Task (Roebers et al., 2011; originally developed by Archibald and Kerns, 1999). Four pages with 25 stimuli each were consecutively presented to the child. Page 1 consisted of colored rectangles (blue, yellow, red, green). Page 2 depicted 4 kinds of fruits and vegetables in appropriate colors (plum = blue, banana = yellow, strawberry = red, lettuce = green). Page 3 presented the same fruits and vegetables but printed in gray. Page 4 again consisted of the same fruits and vegetables, only now they were colored incorrectly. For pages 1 and 2, the children were told to name the color in which items were printed as fast as possible. For pages 3 and 4, children had to name the colors that the fruits/vegetables actually should have (i.e., plum = blue, banana = yellow, etc.). Time (in seconds) needed for naming the colors of all items on each page was measured and an interference score was calculated with higher values indicating more interference: [time p.4 − (time p.1 × time p.3) / (time p.1 + time p.3)] (Archibald and Kerns, 1999).

Hot executive functions. The affective decision-making component of hot EF was measured by an adapted version of the Hungry Donkey Task (Crone and van der Molen, 2004), which is an age-appropriate version of the Iowa Gambling Task, one of the most widely used measures of VM–PFC function (Bechara et al., 1994).

We adapted the task in terms of task duration, instruction, motivational relevance, and complexity, i.e., working memory demands. Four doors (A, B, C, D) were presented side by side on the computer screen (stimulus display; **Figure 1**). Children were told to assist a hungry donkey to collect as many apples as possible by pressing 1 of 4 keys, opening a corresponding door. Moreover, participants were told that they could win a marble if they collected at least 20 apples (in order to enhance motivational relevance). The S, D, K, and L keys of a QWERTZ keyboard were mapped onto the doors from left to right and the left middle, left index, right index, and right middle fingers were assigned to the keys consecutively. Upon pressing one of the keys an outcome display (**Figure 1**) was presented at the position of the opened door, showing the number of (green) apples gained and (red and crossed-out) apples lost in the present trial. Furthermore, the overall sum of gained and lost apples across previous trials was displayed as a positive or negative number below the door. The task consisted of 60 trials. Doors A and B (as well as doors C and D) were identical in their underlying win/loss contingencies. Selecting doors A or B resulted in a gain of 4 apples, whereas selecting doors C or D resulted in a gain of only 2 apples. However, doors A and B were disadvantageous in the long run because after selecting doors A or B 10 times, the participant received 40 apples but had also encountered 5 unpredicted losses of 8, 10, 10, 10, or 12 apples, resulting in a net loss of 10 apples. Choosing doors C or D 10 times, in contrast, resulted in a gain of 20 apples with 5 unpredicted losses of 1, 2, 2, 2, or 3 apples, incurring a net gain of 10 apples. The dependent variable was the net-score difference between advantageous and disadvantageous choices [(C+D)–(A+B)] of the last 50 trials (e.g., Crone et al., 2005). The first 10 trials were excluded from the analysis in order to tap decision making under risk, rather than decision making under ambiguity, because win/loss contingencies have probably not yet been experienced during the first trials (Brand et al., 2007).

To measure the delay of gratification component of hot EF children were asked to choose between receiving a smaller reward immediately or a more valuable reward 1 week later (at the second test session; adapted from Wulfert et al., 2002). In 4 trials (1 vs. 2

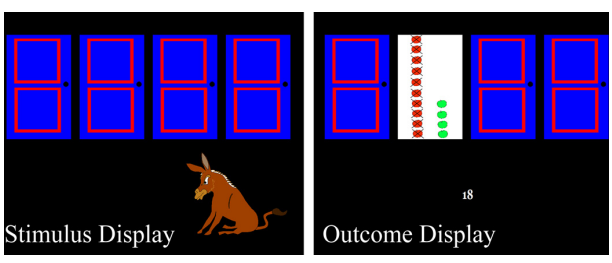


FIGURE 1 | Stimulus and outcome displays of the Hungry Donkey Task (Crone and van der Molen, 2004).

chocolate drops; 1 vs. 5 chewing candies; 1 vs. 2 bouncing frogs; 1 vs. 3 tattoos), the child always saw the immediate but not the delayed reward. The dependent variable was the number of trials in which the child chose to delay.

In a pretest on 41 children (54% females) aged 8–9 years ($M = 8.41$, $SD = 0.49$) the number of delayed trials showed positive associations in the medium range ($r = 0.31$–0.37, $p \leq 0.05$) with academic delay of gratification (Academic Delay of Gratification Scale for Children; Zhang et al., 2011), delay of gratification in eating (subscale from the Delaying Gratification Inventory; Hoerger et al., 2011), and impulsivity (German version of Eysenck's I6 Impulsivity Scale; Stadler et al., 2004), indicating good convergent validity. Furthermore, the four trials were highly associated with a longer version of the task (8 items, $r = 0.88$, $p < 0.001$).

Food approach behavior and weight assessment

Parents rated the food approach behavior of their children on selected items of 4 scales (3 items each; 5-point-response format: 1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = always) of the Children's Eating Behavior Questionnaire (CEBQ; Wardle et al., 2001): Food Responsiveness (e.g., "My child's always asking for food"), Emotional Overeating (e.g., "My child eats more when worried"), Enjoyment of Food (e.g., "My child enjoys eating"), Desire to Drink (e.g., "If given the chance, my child would always be having a drink"), and on the scale External Eating (5 items; e.g., "My child has a desire to eat when s/he watches others eat") of the Dutch Eating Behavior Questionnaire (DEBQ; Van Strien et al., 1986; 4-point scale: 1 = never, 2 = rarely, 3 = sometimes, 4 = often). Furthermore, the children rated their tendency for Restrained Eating (5 items; e.g., "I try to eat less to avoid weight gain"; DEBQ-C; Franzen and Florin, 1997; Van Strien and Oosterveld, 2008; 4-point scale). Although conceptually, restrained eating is not a food approach behavior, it belongs to a category of eating styles leading to higher weight gain in the long-term (Van Strien and Oosterveld, 2008). Therefore, we subsumed it under the term food approach behavior.

Items of the CEBQ and DEBQ were translated into German and back-translated by a native English speaker. Due to time limits, scales were shortened to 3 (CEBQ), 4 (DEBQ: restrained eating) or 5 (DEBQ: external eating) items with those items being selected that displayed the highest factor loadings. However, a broad content spectrum was intended at the same time. The factorial structure of the two questionnaires remained the same, and internal consistency of scales was acceptable to good (Cronbach's α: 0.71-0.89).

Children's body weight was assessed via calibrated digital scales and height was measured using calibrated ultrasound measurement devices, after shoes, hats and heavy jackets had been removed. A standardized BMI score (BMI-SDS; Kromeyer-Hauschild et al., 2001) was calculated in order to ensure comparability across age and gender.

Fluid intelligence

Fluid intelligence was assessed by the Number-Symbol Test of the German version of the Wechsler Intelligence Scale for Children (Petermann and Petermann, 2007).

The child is required to assign symbols to either 5 simple figures (for ages 6–7 years, version A) or to 9 digits (for ages 8–16 years, version B) as quickly as possible. For both versions A and B, the dependent variable is the amount of correct symbols allocated within 120 s (standardized T-Values were calculated).

PROCEDURE

Measures were administered as part of a multifaceted study on intrapersonal developmental risk factors in childhood (February–December 2012). Children completed two 50-min assessments with an interval of about 7 days, conducted by trained and supervised doctoral students or research assistants. Each child was tested individually by one experimenter during the morning hours in a quiet room either at school or at home. Tasks were performed in a counterbalanced order (Blocks of ABCD/BADC). Subsequent analyses, however, revealed no effect of task sequence.

Parents answered the eating behavior questionnaires either online or in printed format. Questionnaires were mostly answered by mothers (71%) or both parents (21%). All participants were guaranteed privacy and children received a cinema voucher as reward upon completion.

All procedures were approved by the Research Ethics Board at the University of Potsdam and by the Ministry of Education, Youth and Sport of the Federal State of Brandenburg. Children and parents were informed about the procedure, materials, and study aims prior to their participation. For each child, informed consent was obtained from a primary caregiver.

STATISTICAL ANALYSES

Research questions were answered using structural equation modeling (SEM). Models were fit using MPlus Version 7.11 (Muthén and Muthén, 1998–2012). For the first research question we conducted a confirmatory factor analysis (CFA). A one-factor and a two-factor model were fit to the 5 EF tasks. The one-factor model postulated that the tasks can be best described by a unitary higher-level construct. The two-factor model assumed that the tasks can be best conceptualized by a hot and a cool dimension, which are dissociable but correlated. In order to compare the one-factor and the two-factor model on a descriptive level, the model with the lowest Akaike Information Criterion (AIC) was regarded as the best fitting model (Schermelleh-Engel et al., 2003). Furthermore, in order to examine whether the measurement model differs between the younger (<8.4 years) and the older (>8.4 years) half of the sample (median split) we used a χ^2 difference test to compare a CFA model that estimated factor loadings freely to a CFA model that constrained factor loadings to be equal across groups. The second research question was examined using a SEM in which the 6 eating behavior scales were entered as latent variables and regressed on hot and cool EF, grouping children by gender. Age and fluid intelligence, which are known to be related to EF and eating behavior, were controlled for.

Model fit was evaluated using a combination of absolute (standardized root mean residual, SRMR; root mean squared error of approximation, RMSEA) and comparative (comparative fit index, CFI) fit indices. Model fit was considered good if CFI ≥ 0.97, RMSEA ≤ 0.05 and SRMR ≤ 0.05 (Schermelleh-Engel et al.,

2003). We did not rely upon the χ^2 statistic to evaluate model fit because the value of p associated with the χ^2 statistic is related to sample size and was therefore considered to be overly sensitive to misfits (Schermelleh-Engel et al., 2003). An alpha level of $p \leq 0.05$ was used for all statistical tests.

The percentage of missing values was ≤1.3 for the child-assessed data (EF, BMI) and ≤ 19.6 for the parent-assessed data (Food Approach). Assuming data to be missing at random we estimated missing values by full information maximum likelihood (FIML) estimation. Results, however, did not differ when analyzing complete cases only.

RESULTS

DESCRIPTIVE STATISTICS AND INTERCORRELATIONS

Bivariate correlations, as well as means and standard deviations, for all of the variables included in structural equation models are summarized in **Table 1**. On average, children were quite good at solving the attention shifting and the delay of gratification task and they showed medium scores on updating, inhibition, and affective decision-making. The performance on different EF tasks was positively, albeit low to modestly, intercorrelated (Variables 1–5). The 3 cool EF tasks showed low to modest positive correlations with the fluid intelligence measure, whereas the 2 hot EF tasks did not. Furthermore, performance on all EF tasks was positively associated with age. On average, boys showed slightly better performance on the 2 hot EF tasks than did girls, whereas girls outperformed boys in the cool EF inhibition and attention shifting tasks.

Children showed low to medium levels of food approach behavior with the highest scores on enjoyment of food. This is broadly in line with results from other studies (Sleddens et al., 2008; Van Strien and Oosterveld, 2008). All food-approach scales were positively correlated with one another and with BMI-SDS, mostly to a medium extent. Girls scored a bit higher than boys on the food responsiveness scale; no other gender differences were apparent.

FACTOR STRUCTURE OF EF TASKS

One-factor (general) vs. two-factor (hot/cool) EF model in the overall sample

The first aim of this study was to examine whether EF measures can be divided into a hot and cool factor or whether they actually correspond to a unitary construct in middle childhood. A one-factor CFA model fitted the data well, $\chi^2(5) = 3.54$, $p = 0.62$, $CFI = 1.00$, $RMSEA = 0.00$, $SRMR = 0.01$, AIC = 43,773.59. Standardized parameter estimates are provided in **Figure 2**. The 3 cool EF tasks showed similar medium-sized factor loadings. However, the standardized factor loadings of the 2 hot EF tasks were significant but very low, falling under a general cutoff value (0.40) for the inclusion into one factor (Stevens, 2001).

A two-factor CFA model also fitted the data well, $\chi^2(4) = 3.34$, $p = 0.50$, $CFI = 1.00$, $RMSEA = 0.00$, $SRMR = 0.01$, $AIC = 44{,}775.39$. Standardized factor loadings indicated that all 3 cool EF tasks made a nearly equally strong contribution to the cool EF latent variable. However, standardized factor loadings of the 2 hot EF tasks were again very low and only marginally significant indicating that the tasks were not represented well by one underlying hot EF factor. Moreover, there was a high positive correlation between the hot and cool EF latent factors (**Figure 3**).

Table 1 | Descriptive statistics and intercorrelations of assessed variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Attention shifting															
2. Updating	0.33**														
3. Inhibition[a]	−0.33**	−0.27**													
4. Decision making	0.10**	0.05*	−0.03												
5. Delay of gratification	0.09**	0.08**	−0.06*	0.03											
6. Enjoyment of food	0.01	−0.02	−0.05	−0.02	−0.02										
7. Desire to drink	−0.10**	−0.06*	0.03	−0.02	0.03	0.05									
8. Food responsiveness	−0.03	−0.03	−0.01	−0.00	−0.03	0.44**	0.26**								
9. Emotional overeat	0.01	−0.04	0.00	−0.03	0.01	0.10**	0.24**	0.39**							
10. External eating	−0.01	−0.01	0.00	−0.02	0.02	0.41**	0.16**	0.48**	0.27**						
11. Restrained eating	−0.06*	−0.04	0.03	0.02	0.09	0.08**	0.09**	0.20**	0.08**	0.05					
12. BMI-SDS	−0.10**	−0.06*	0.04	−0.02	−0.01	0.24**	0.19**	0.44**	0.17**	0.17**	0.30**				
13. Male[b]	−0.18**	0.02	0.10**	0.13*	0.09**	−0.06	0.03	−0.09**	0.03	−0.03	0.02	0.00			
14. Age	0.27**	0.22*	−0.34**	0.06*	0.11*	0.02	−0.01	0.05	0.01	−0.01	0.01	−0.00	0.04		
15. Fluid intelligence[c]	0.15**	0.10**	−0.26**	−0.00	0.02	0.07*	0.02	0.04	0.00	−0.01	−0.03	−0.05*	−0.22**	−0.07**	
Mean	15.6	6.1	24.9	5.5	2.8	3.5	1.9	1.6	1.2	2.9	2.2	0.16	/	8.4	51.4
Standard deviation	4.7	1.6	8.8	11.4	1.2	0.83	0.93	0.89	0.41	0.57	0.84	1	/	0.95	9.2
Min-Max (theoretical)	0–22	0–16	0–89[d]	−60–60	0–4	1–5	1–5	1–5	1–5	1–4	1–4	−4.2–3.6[d]	/	6–11	27–80[d]

N = 1657. Variables 1–5 are indicators of cool (1–3) and hot (4–5) EF; variables 6–11 are facets of food approach behavior.

[a] Interference measure (negatively polarized); [b] Value labels: 1 = male, 0 = female; [c] T-Value Number-Symbol-Test; [d] Min and/or Max values are theoretically infinite, thus table values are sample-specific.

**$p \leq 0.05$; **$p \leq 0.01$.*

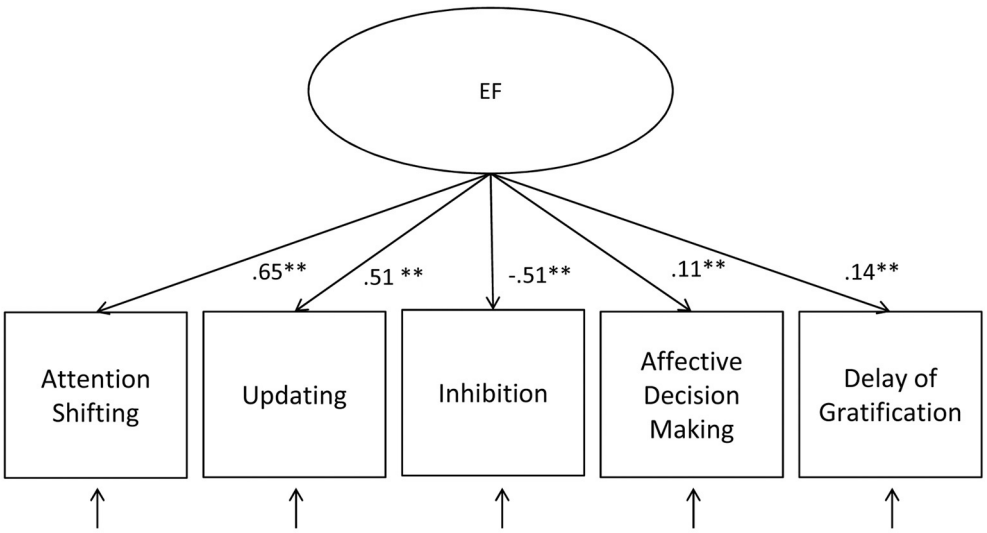


FIGURE 2 | One-factor CFA model of EF tasks. $^{**}p \leq 0.01$.

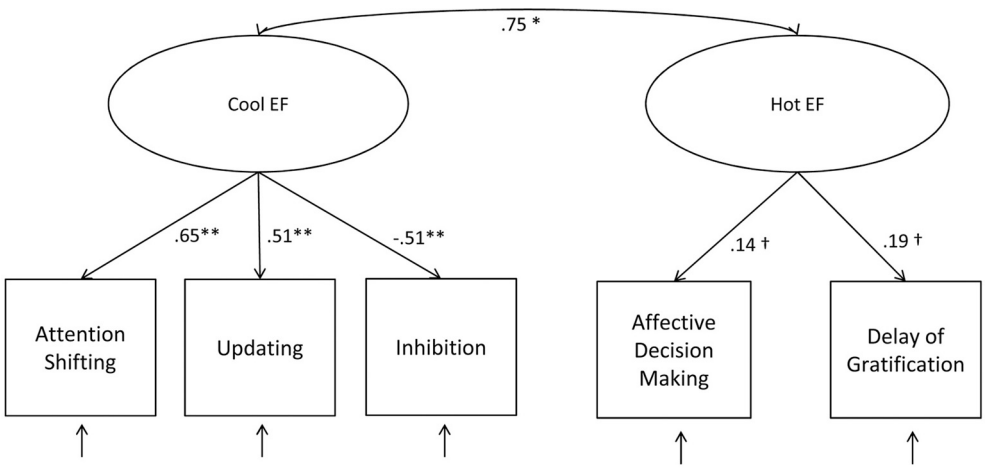


FIGURE 3 | Two-factor CFA model of EF tasks. $^{\dagger}p \leq 0.10$; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

When comparing models on a descriptive level, the slightly lower AIC value indicated that the one-factor model seemed to be a better tradeoff between model fit and model complexity than the two-factor model.

One-factor (general) vs. two-factor (hot/cool) EF model in younger and older children

As a second step, we examined whether the measurement model differed between the younger (<8.4 years) and the older (>8.4 years) half of the sample (median split). Standardized factor loadings for the one-factor and the two-factor model within both age groups are shown in **Table 2**.

First, the one-factor CFA model was tested in the younger and older age group separately revealing a good fit within both age groups: Younger children: $\chi^2\,(5) = 3.50$, $p = 0.62$, $CFI = 1.00$, $RMSEA = 0.00$, $SRMR = 0.01$; Older children: $\chi^2\,(5) = 1.63$, $p = 0.90$, $CFI = 1.00$, $RMSEA = 0.00$, $SRMR = 0.01$.

Then, a one-factor CFA model that estimated factor loadings freely (A) was tested against a one-factor CFA model that constrained factor loadings to be equal across groups of younger and older children (B) in order to determine whether model fit worsened significantly. Intercepts were constrained to be equal in both models. Both models fitted the data well: Model (A): $\chi^2\,(14) = 24.15$, $p = 0.04$, $CFI = 0.97$, $RMSEA = 0.03$, $SRMR = 0.03$; Model (B): $\chi^2\,(18) = 27.09$, $p = 0.08$, $CFI = 0.97$, $RMSEA = 0.03$, $SRMR = 0.03$. A χ^2 difference test revealed no significant worsening of fit of the constrained model as compared to the free model, $\Delta\chi^2\,(4) = 2.95, p = 0.57$, suggesting that factor-loadings were equal across groups of younger and older children. In this instance, hot EF loaded very low on the general EF factor in both age groups.

Table 2 | Standardized factor loadings for EF tasks on the one- and two-factor model in the younger (<8.4 years) and older (>8.4 years) half of the sample.

	1-factor-model		2-factor-model			
	General-EF factor		Hot-EF factor		Cool-EF factor	
	Younger[b]	Older[c]	Younger[b,d]	Older[c]	Younger[b,d]	Older[c]
Attention shifting	0.58**	0.49**			/	0.66**
Updating	0.50**	0.38**			/	0.44**
Inhibition[a]	−0.43**	−0.61**			/	−0.45**
Decision making	0.11*	0.11**	/	0.12		
Delay of gratification	0.12*	0.13**	/	0.19		

[a] *Interference measure (negatively polarized);* [b] *younger half of the sample;* [c] *older half of the sample;* [d] *two-factor model could not be estimated for the younger half of the sample.*

*$p \leq 0.05$; **$p \leq 0.01$.

In a second step, a two-factor CFA model was tested separately within both age groups. The two-factor model fitted the data well in the older subgroup, $\chi^2\,(4) = 1.52, p = 0.82$, $CFI = 1.000$, $RMSEA = 0.000$, $SRMR = 0.009$, with hot EF factor loadings being again low and not significant. However, it was not possible to estimate the two-factor model in the subgroup of younger children, seemingly due to the absent covariance of hot EF tasks. This suggests that the proposed two-factor model was highly inconsistent with the data, implying that a differentiation of EF into a hot and a cool component does not seem plausible for children aged between 6.0 and 8.4 years in our sample.

ASSOCIATIONS BETWEEN EF AND FOOD APPROACH BEHAVIOR

The second aim was to examine how hot and cool EF are associated with different eating styles that put children at risk of becoming overweight. We expected that a difficulty in self-control, seen in lower performance in hot and cool EF tasks, would go along with higher occurrence of food approach behavior.

As the 2 hot EF tasks did neither load well onto one hot EF factor, nor onto the general EF factor, they were further analyzed separately as 2 manifest variables. In contrast, the 3 cool EF tasks were entered as one latent cool EF factor. A SEM was estimated, in which the ratings of children's food approach behavior (6 scales) were regressed on the latent cool EF factor as well as on the 2 manifest hot EF variables, including age (as continuous variable) and fluid intelligence as covariates. Standardized parameter estimates for significant associations ($p \leq 0.05$) are reported in **Figure 4**. The SEM fitted the data well, $\chi^2\,(632) = 1058.31$, $p < 0.01$, $CFI = 0.97$, $RMSEA = 0.03$, $SRMR = 0.04$. In girls, cool EF showed relatively small but significant associations with 3 out of 6 eating styles, namely desire to drink, food responsiveness, and restrained eating. Furthermore, the hot EF component delay of gratification was slightly positively related to emotional overeating. However, there were neither significant associations of cool EF and emotional overeating, enjoyment of food, or external eating, nor of the hot EF component affective decision-making and any of the eating styles. In boys, neither hot nor cool EF were significantly associated with any aspect of food approach behavior.

However, using a chi-square difference test to examine differences in regression coefficients between boys and girls revealed a significant moderation by gender only for the association between cool EF and restrained eating, $\Delta\chi^2\,(1) = 4.74$, $p = 0.03$, and between delay of gratification and emotional overeating, $\Delta\chi^2\,(1) = 5.15$, $p = 0.02$. Neither the association between cool EF and desire to drink, $\Delta\chi^2\,(1) = 0.16$, $p = 0.69$, nor between cool EF and food responsiveness, $\Delta\chi^2\,(1) = 0.23$, $p = 0.63$, differed significantly between boys and girls.

DISCUSSION

STRUCTURE OF HOT AND COOL EF IN MIDDLE CHILDHOOD

To date, findings on the structure of hot and cool EF in children have been inconsistent and mostly based on preschool samples. The first aim of this study was to investigate whether performance on EF tasks can be distinguished into correlated hot and cool components (two-factor model) or whether it is better represented by one general EF-factor (one-factor model) in a large sample of children from German school classes 1–3 (aged 6–11 years).

Our data shows that the one-factor as well as the two-factor model fit the data well, with information parameter indices denoting the one-factor EF model to be a better tradeoff between model fit and model complexity. However, standardized factor loadings of the hot EF tasks (affective decision-making, delay of gratification) were very low on both models, falling under a general cutoff value (0.40) for the inclusion into one factor (Stevens, 2001). The cool EF tasks on the other hand showed similar, good factor loadings in both models indicating that the cool components of EF, that is, attention shifting, inhibition, and updating, are highly associated in middle childhood. Comparing subgroups of younger (<8.4 year-olds) and older (>8.4 year-olds) children in our sample, the one-factor model applied to both age groups with equal factor loadings within samples. However, for the younger subgroup only the one-factor-model fit the data well. A two-factor-model could not be estimated due to the missing covariance between the hot EF tasks.

Findings suggest that whereas cool EF seems to be a coherent functional construct in middle childhood, hot EF does not. The two hot EF tasks were neither represented well by the one-factor nor by the two-factor model, which points to the possibility that hot EF is a more complex and heterogeneous construct than

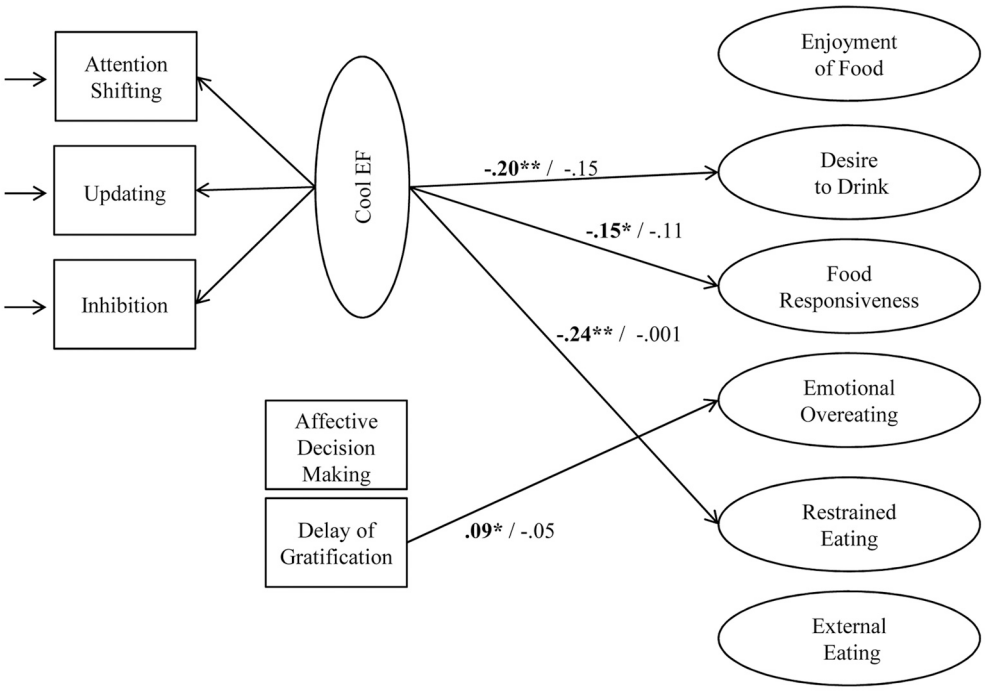


FIGURE 4 | Associations of the latent cool EF factor and the two hot EF tasks with latent food approach scales. Results are controlled for age and fluid intelligence. Results of girls are highlighted in bold, results of boys are shown in regular font. $^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

originally thought. The minor loadings on the one-factor model indicate that the hot EF tasks did not share a large amount of variance with cool EF, supporting the idea of different underlying mechanisms between hot and cool EF (Zelazo and Müller, 2002). This is further confirmed by differential relations of hot and cool EF to fluid intelligence and gender. Cool but not hot EF tasks were related to a fluid intelligence measure, which has also been proposed previously (e.g., Bar-On et al., 2003; Hongwanishkul et al., 2005). Furthermore, on average, girls outperformed boys in the cool inhibition and attention shifting tasks, whereas boys showed slightly better performance than girls on the two hot EF tasks. The latter is in line with results showing that men outperform women on the Iowa Gambling Task (Reavis and Overman, 2001) and with studies suggesting that VM–PFC develops more rapidly in males than in females (e.g., Clark and Goldman-Rakic, 1989; Overman et al., 1996).

At the same time the minor factor loadings of hot regulatory tasks on a single hot EF factor reflect their missing covariance indicating that hot EF is not a particular homogeneous construct in itself. This confirms negative evidence for a single hot EF factor in younger children, suggesting that the construct of hot EF may need to be further refined (Hongwanishkul et al., 2005; Prencipe et al., 2011). This seems to contradict studies that found substantial correlations within hot EF tasks (e.g., Sonuga-Barke et al., 2003; Smith-Donald et al., 2007; Brock et al., 2009; Willoughby et al., 2011). However, those studies all used variants of delay of gratification tasks to assess hot EF (e.g., snack delay, toy wrap, tongue task) requiring children to wait and inhibit themselves in order to get a reward. In contrast, the two hot EF tasks used in the present study were conceptually less similar, and other studies using variants of those tasks also failed to find evidence for a single hot EF factor. Hongwanishkul et al. (2005) even reported a negative association between the Children's Gambling Task and a delay of gratification task in 3- to 5-year-olds. Similarly, Prencipe et al. (2011) found the Children's Gambling Task not to be associated with a delay discounting task and, consistent with the present results, both tasks loaded only marginally onto a single EF factor in 8- to 11-year-olds. However, in cocaine-dependent adults, the Iowa Gambling Task showed positive relations to delay discounting (Monterosso et al., 2001).

The missing covariance between delay of gratification and affective decision-making can be explained by some fundamental task differences. For instance, both tasks differ considerably in their working memory demands (Hongwanishkul et al., 2005). Whereas the present affective decision-making task required tracking wins and losses across a series of 60 trials, the delay of gratification task involved only 4 independent choices. Furthermore, the two hot EF tasks differed with respect to the time that children had to wait for the rewards and to the certainty with which rewards were obtained. Whereas choice contingencies are clear in the delay of gratification task (one now vs. more 1 week later), they remain purposely unclear in the affective decision-making task (Hongwanishkul et al., 2005).

Moreover, because EF does not develop in a homogenous fashion (e.g., Passler et al., 1985), affective decision-making and delay of gratification, although conceptually related, may evolve at different time points. This would make developmental covariance less likely and is suggested by diverging task difficulties. Whereas

the delay of gratification task used in the present study is a rather simple measure of hot EF, the affective decision-making task is relatively complex, probably also placing stronger demands on non-executive skills. This was also reflected by the sample distribution because the delay of gratification task showed some ceiling effects. In contrast, the affective decision-making task proved more difficult to be completed successfully.

Altogether, our results support a distinction between hot and cool facets of EF, but further investigation is needed in order to examine whether hot EF may itself be a heterogeneous construct. This has also been suggested by other authors examining younger and older populations of children (Hongwanishkul et al., 2005; Prencipe et al., 2011). Furthermore, our results on developmental changes in the structure of EF show that whereas performance on all EF tasks was positively associated with age, there was no significant developmental change in the covariance between tasks, disagreeing with the hypothesized idea of a growing differentiation between hot and cool EF for a population of 1st to 3rd graders (Johnson, 2011; Zelazo and Carlson, 2012).

EF AND FOOD APPROACH BEHAVIOR

The second aim of the present study was to examine whether EF performance is associated with food approach behavior in a population sample of 6- to 11-year-old children. Whereas much is known about EF deficits in clinical populations of the overweight and obese (Smith et al., 2011) there is very little information on how hot and cool EF are associated with eating styles that put children at risk for the development of overweight.

The present study revealed expected negative associations between EF and several food-approach behaviors for girls, but not for boys. After controlling for age and fluid intelligence, in girls lower cool EF went together with higher scores on 3 out of 6 food-approach scales, namely food responsiveness, desire to drink, and restrained eating. No significant associations occurred between cool EF and enjoyment of food, emotional eating, or external eating. Unexpectedly, as for hot EF, performance in the delay of gratification task showed a small *positive* relation to emotional overeating, and the affective decision-making task was not at all associated with food-approach behavior. In boys, neither hot nor cool EF were associated with any of the eating styles. However, the difference in regression coefficients between boys and girls was only significant for the associations between cool EF and restrained eating and between affective decision-making and emotional overeating. All significant regression coefficients were in the low to medium range (Cohen, 1988). However, when interpreting the strength of the associations, the different assessment methods (EF: child experiments vs. food approach: mostly parent ratings) have to be kept in mind (Campbell and Fiske, 1959).

Results show that lower EF cannot only be found in overweight or obese individuals (Smith et al., 2011) but that EF is linearly associated with food-approach styles that are presumed to be a risk factor for the development of overweight in a normal population of children. This is in line with findings that increased dysexecutive traits are associated with disinhibited eating and greater food cravings in a population sample of adults (Spinella and Lyke, 2004). Thus, EF plays a role in eating, even in normal populations of children, suggesting that eating disorders or obesity represent only the extremes of a normal continuum of eating behavior. This is contrary to the assumption of Smith et al. (2011) that negative effects of adiposity on cognition might be only detected in populations who exceed a threshold, i.e., only in the obese.

Although neuroimaging studies suggest that the VM-PFC, which is associated with hot EF, plays a role in the reinforcing value of food, satiety, and the control of eating (Rolls, 2004), we did not find the expected negative associations between hot EF and food approach behavior. However, the facets of hot regulation assessed in our study may not be that relevant for the regulation of eating in normal populations. This might be especially true for the affective decision-making task because its relation to eating on the behavior level is quite subtle.

Furthermore, to date there is only little information on developmental correlates of hot EF. It can be speculated that hot EF will take effect only when severely impaired or over a longer period of time. It might also show its impact on eating only later in development as affective decision-making is believed to develop quite late, with adult levels not being reached until late adolescence (Crone and van der Molen, 2004). This is also reflected in relatively low performance levels in the present sample. Moreover, performance on hot EF tasks may not only result from an inability or cognitive dysfunction but also from unwillingness or a motivational dysfunction (Reynolds and Schiffbauer, 2005; Willoughby et al., 2011), which might bias associations with other variables.

However, we found a low positive association between delay of gratification and emotional overeating. This is surprising and seems to contradict findings showing that obese children have greater difficulty waiting for a larger, delayed reward than children of normal weight (Johnson et al., 1978; Bonato and Boland, 1983). Yet, other authors failed to find such differences between overweight and normal-weight children (Geller et al., 1981; Bourget and White, 1984). However, in a normal child population eating in response to negative emotions might rather be a maladaptive strategy of emotion regulation than an act of impulsivity. Being able to resist a reward requires affect regulation as well, what might explain the small positive association between delay of gratification and emotional overeating. However, this is only speculative and considering the low effect size, this association should not be overrated.

We found associations between cool EF and 3 out of 6 food-approach styles in girls, namely food responsiveness, desire to drink, and restrained eating. Yet, no significant associations occurred between cool EF and enjoyment of food, emotional eating, or external eating. Food responsiveness and desire to drink refer to eating styles that imply a constant need for food or drink. Moreover, restrained eating is often initiated as a response to weight gain (Johnson et al., 2012) probably in order to compensate for lower EF. Thus, those 3 eating styles that were associated with EF might be the more obvious signs of a lack of self-regulation ability as compared to the others.

Associations between EF and food-approach behavior were only found in girls, but not in boys, suggesting that self-regulatory abilities do not play a role in food-approach behavior of elementary school-aged boys. At this age, boys probably self-regulate their own eating behavior less than girls do. This might be

due to the facts that in Western cultures the pressure to be thin is much higher for girls and women than for men, and that women face more stringent standards of physical appearance (Friedman and Reichmann, 2002). Women also report more weight stigmatization, starting at lesser degrees of being overweight (e.g., Cossrow et al., 2001) and they suffer more from being overweight than men (Van der Merwe, 2007). Gender differences have also been reported by studies assessing covariates of overweight. For instance, obese girls, but not obese boys, suffer in their ability to focus attention (Mond et al., 2007), suggesting gender-specific associations between obesity and impairments in specific aspects of developmental functioning. Moreover, cross-sectional as well as longitudinal studies found a moderating role of gender in the association between personality factors and body weight (e.g., Brummett et al., 2006; Armon et al., 2013). Positive relations between neuroticism and body weight, and negative relations between conscientiousness and body weight were found to be stronger for women than for men. Likewise, openness was negatively associated with body weight for women, but not for men. Although the present findings did not reveal any relations between EF and food-approach behavior in boys, these might just occur at a later age, as soon as pubertal development makes dealing with body-weight issues and the resulting conscious regulation of eating more relevant for boys.

One limitation of the present study is that children's performance in hot and cool EF was measured by only 5 tasks. Future studies would benefit from the inclusion of a greater number of indicators, especially for hot EF in order to further examine possible differences within hot EF. Moreover, although the hot EF tasks were mainly not associated with food approach behavior in our sample, this does not imply that hot EF is less important than cool EF for the regulation of eating. Probably, applying ecologically more relevant tasks could have helped to show associations of hot EF with eating behavior.

CONCLUSION

The present study examined the structure of hot and cool EF and its relation to food approach behavior in a representative sample of elementary-school children from school class 1–3. Results showed that cool EF seems to be a reasonable coherent functional construct in middle childhood. However, further clarification is required regarding the construct of hot EF. Nevertheless, hot and cool EF do not seem to share exactly the same underlying mechanisms, and their distinction is supported by differential relations to fluid intelligence and food-approach behavior, as well as by gender differences in task performance. Therefore, as has been noted by other authors (e.g., Hongwanishkul et al., 2005), it needs to be further examined to what extent hot EF—although distinct from cool EF—might not be a homogeneous construct itself.

Furthermore, the study provides first evidence that not only obesity is associated with impaired EF, but that linear associations between hot and cool EF and the occurrence of food approach behaviors occur in a normal population of elementary school-aged girls. This extends findings on relationships of prefrontal neural systems and eating from clinical populations, e.g., patients showing neurological or eating disorders (e.g., Dempsey et al., 2011; Smith et al., 2011), into the normal population. Considering these results, it seems plausible to assume that EF constitutes a risk factor for eating styles that contribute to the development of overweight. However, results of the present study rely on cross-sectional data. Longitudinal designs examining relations between earlier EF and later eating behavior are needed to shed light on the important question of whether EF is a risk factor for the development of obesity or whether in turn the type of diet is responsible for cognitive deficits (Smith et al., 2011).

Today's oversupply of palatable high-caloric food is known to play an important role in promoting obesity (Hill and Peters, 1998) but not all individuals exposed to this environment become overweight or obese. Determining modifiable risk factors of obesity is of particular importance given that obesity is currently considered one of the most increasingly important health issues (WHO, 2006; Moß et al., 2007). As there is evidence that EF capacity can be improved (e.g., Klingberg et al., 2005; Diamond and Barnett, 2007) and that EF improvement helps patients suffering from eating disorders (Tchanturia et al., 2007; Genders et al., 2008), the training of EF appears to be a promising tool for the prevention of overweight and obesity in children. Thus, examining the exact role of EF for the development of obesity seems to be an important topic for future research.

AUTHOR CONTRIBUTIONS

All authors have contributed considerably to the conception and design of the work including the formulation of hypotheses. Karoline Groppe has primarily analyzed and interpreted the data, whereas Birgit Elsner has formulated the problem and revised the work. Both authors have agreed to be accountable for all aspects of the work and to submit the manuscript in this form.

ACKNOWLEDGMENTS

The authors would like to thank Sebastian Grümer for his helpful methodological suggestion and Stefanie Thies for her substantial help in data collection. This research was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG; GRK 1668/1).

REFERENCES

Allan, N. P., and Lonigan, C. J. (2011). Examining the dimensionality of effortful control in preschool children and its relation to academic and socioemotional indicators. *Dev. Psychol.* 47, 905–915. doi: 10.1037/a0023748

Alvarez, J. A., and Emory, E. (2006). Executive function and the frontal lobes: a meta-analytic review. *Neuropsychol. Rev.* 16, 17–42. doi: 10.1007/s11065-006-9002-x

Appelhans, B. M. (2009). Neurobehavioral inhibition of reward-driven feeding: implications for dieting and obesity. *Obesity* 17, 640–647. doi: 10.1038/oby.2008.638

Archibald, S. J., and Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychol.* 5, 115–129. doi: 10.1076/chin.5.2.115.3167

Armon, G., Melamed, S., Shirom, A., Shapira, I., and Berliner, S. (2013). Personality traits and body weight measures: concurrent and across-time associations. *Eur. J. Pers.* 27, 398–408. doi: 10.1002/per.1902

Bar-On, R., Tranel, D., Denburg, N. L., and Bechara, A. (2003). Exploring the neurological substrate of emotional and social intelligence. *Brain* 126, 1790–1800. doi: 10.1093/brain/awg177

Batterink, L., Yokum, S., and Stice, E. (2010). Body mass correlates inversely with inhibitory control in response to food among adolescent girls: An fMRI study. *Neuroimage* 52, 1696–1703. doi: 10.1016/j.neuroimage.2010.05.059

Bechara, A. (2004). The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. *Brain Cogn.* 55, 30–40. doi: 10.1016/j.bandc.2003.04.001

Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994). Insensitivity to future consequences following damage to the human prefrontal cortex. *Cognition* 50, 7–15. doi: 10.1016/0010-0277(94)90018-3

Bekker, M. H. J., van de Meerendonk, C., and Mollerus, J. (2004). Effects of negative mood induction and impulsivity on self-perceived emotional eating. *Int. J. Eat. Disord.* 36, 461–469. doi: 10.1002/eat.20041

Bonato, D. P., and Boland, F. J. (1983). Delay of gratification in obese children. *Addict. Behav.* 8, 71–74. doi: 10.1016/0306-4603(83)90059-X

Bourget, V., and White, D. R. (1984). Performance of overweight and normal-weight girls on delay of gratification tasks. *Int. J. Eat. Disord.* 3, 63–71.

Brand, M., Recknor, E. C., Grabenhorst, F., and Bechara, A. (2007). Decisions under ambiguity and decisions under risk: correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *J. Clin. Exp. Neuropsychol.* 29, 86–99. doi: 10.1080/13803390500507196

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., and Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Child. Res. Q.* 24, 337–349. doi: 10.1016/j.ecresq.2009.06.001

Brummett, B. H., Babyak, M. A., Williams, R. B., Barefoot, J. C., Costa, P. T., and Siegler, I. C. (2006). NEO personality domains and gender predict levels and trends in body mass index over 14 years during midlife. *J. Res. Pers.* 40, 222–236. doi: 10.1016/j.jrp.2004.12.002

Bryant, E. J., King, N. A., and Blundell, J. E. (2008). Disinhibition: its effects on appetite and weight regulation. *Obes. Rev.* 9, 409–419. doi: 10.1111/j.1467-789X.2007.00426.x

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56, 81–105. doi: 10.1037/h0046016

Clark, A. S., and Goldman-Rakic, P. S. (1989). Gonadal hormones influence the emergence of cortical function in nonhuman primates. *Behav. Neurosci.* 103, 1287–1295. doi: 10.1037/0735-7044.103.6.1287

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences, 2nd Edn.* New Jersey, NJ: Erlbaum.

Cohen, J. I., Yates, K. F., Duong, M., and Convit, A. (2011). Obesity, orbitofrontal structure and function are associated with food choice: a cross-sectional study. *BMJ open* 1:e000175. doi: 10.1136/bmjopen-2011-000175

Cortese, S., Angriman, M., Maffeis, C., Isnard, P., Konofal, E., Lecendreux, M. et al. (2008). Attention-deficit/hyperactivity disorder (ADHD) and obesity: a systematic review of the literature. *Crit. Rev. Food Sci. Nutr.* 48, 524–537. doi: 10.1080/10408390701540124

Cossrow, N. H., Jeffery, R. W., and McGuire, M. T. (2001). Understanding weight discrimination: a focus group study. *J. Nutr. Educ.* 33, 208–214. doi: 10.1016/S1499-4046(06)60033-X

Crone, E. A., Bunge, S. A., Latenstein, H., and van der Molen, M. W. (2005). Characterization of children's decision making: sensitivity to punishment frequency, not task complexity. *Child Neuropsychol.* 11, 245–263. doi: 10.1080/092970490911261

Crone, E. A., and van der Molen, M. W. (2004). Developmental changes in real life decision making: performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Dev. Neuropsychol.* 25, 251–279. doi: 10.1207/s15326942dn2503_2

Dalen, L., Sonuga-Barke, E. J. S., Hall, M., and Remington, B. (2004). Inhibitory deficits, delay aversion and preschool AD/HD: implications for the dual pathway model. *Neural Plast.* 11, 1–11. doi: 10.1155/NP.2004.1

Dempsey, A., Dyehouse, J., and Schafer, J. (2011). The relationship between executive function, AD/HD, overeating, and obesity. *West. J. Nurs. Res.* 33, 609–629. doi: 10.1177/0193945910382533

Diamond, A., and Barnett, W. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Elfhag, K., and Morey, L. C. (2008). Personality traits and eating behavior in the obese: poor self-control in emotional and external eating but personality assets in restrained eating. *Eat. Behav.* 9, 285–293. doi: 10.1016/j.eatbeh.2007.10.003

Eslinger, P. J., Flaherty-Craig, C. V., and Benton, A. L. (2004). Developmental outcomes after early prefrontal cortex damage. *Brain Cogn.* 55, 84–103. doi: 10.1016/S0278-2626(03)00281-1

Fischer, S., Smith, G. T., and Cyders, M. A. (2008). Another look at impulsivity: a meta-analytic review comparing specific dispositions to rash action in their relationship to bulimic symptoms. *Clin. Psychol. Rev.* 28, 1413–1425. doi: 10.1016/j.cpr.2008.09.001

Franzen, S., and Florin, I. (1997). Der Dutch eating behavior questionnaire für kinder (DEBQ-K) – Ein Fragebogen zur Erfassung gezügelten Essverhaltens [The Dutch Eating behavior questionnaire for children (DEBC-C) – A questionnaire for the assessment of restrained eating]. *Kindheit und Entwicklung* 6, 116–122.

Friedman, K., and Reichmann, S. (2002). Body image partially mediates the relationship between obesity and psychological distress. *Obesity* 10, 33–41. doi: 10.1038/oby.2002.5

Geller, S. E., Keane, T. M., and Scheirer, C. J. (1981). Delay of gratification, locus of control, and eating patterns in obese and nonobese children. *Addict. Behav.* 6, 9–14. doi: 10.1016/S0306-4603(81)80002-0

Genders, R., Davies, H., St Louis, L., Kyriacou, O., Hambrook, D., and Tchanturia, K. (2008). Long-term benefits of CRT for anorexia. *Br. J. Healthc. Manage.* 14, 105–109.

Happaney, K., Zelazo, P. D., and Stuss, D. T. (2004). Development of orbitofrontal function: current themes and future directions. *Brain Cogn.* 55, 1–10. doi: 10.1016/j.bandc.2004.01.001

Hill, J. O., and Peters, J. C. (1998). Environmental contributions to the obesity epidemic. *Science* 280, 1371–1374. doi: 10.1126/science.280.5368.1371

Hoerger, M., Quirk, S. W., and Weed, N. C. (2011). Development and validation of the delaying gratification inventory. *Psychol. Assess.* 23, 725–738. doi: 10.1037/a0023286

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., and Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: age-related changes and individual differences. *Dev. Neuropsychol.* 28, 617–644. doi: 10.1207/s15326942dn2802_4

Jasinska, A. J., Yasuda, M., Burant, C. F., Gregor, N., Khatri, S., Sweet, M., et al. (2012). Impulsivity and inhibitory control deficits are associated with unhealthy eating in young adults. *Appetite* 59, 738–747. doi: 10.1016/j.appet.2012.08.001

Johnson, F., Pratt, M., and Wardle, J. (2012). Dietary restraint and self-regulation in eating behavior. *Int. J. Obes.* 36, 665–674. doi: 10.1038/ijo.2011.156

Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Dev. Cogn. Neurosci.* 1, 7–21. doi: 10.1016/j.dcn.2010.07.003

Johnson, W. G., Parry, W., and Drabman, R. S. (1978). The performance of obese and normal size children on a delay of gratification task. *Addict. Behav.* 3, 205–208. doi: 10.1016/0306-4603(78)90020-5

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., and Gustafsson, P., Dahlström, K. et al. (2005). Computerized training of working memory in children with ADHD-A randomized, controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* 44, 177–186. doi: 10.1097/00004583-200502000-00010

Kromeyer-Hauschild, K., Wabitsch, M., Kunze, D., Geller, F., and Geiß, H. C., Hesse, V., et al. (2001). Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben [Percentiles for the body mass index in childhood and adolescents using diverse German samples]. *Monatsschrift Kinderheilkd.* 149, 807–818. doi: 10.1007/s001120170107

Kurth, B.-M., and Schaffrath Rosario, A. (2007). The prevalence of overweight and obese children and adolescents living in Germany. Results of the german health interview and examination survey for children and adolescents (KiGGS). *Bundesgesundheitsblatt* 50, 736–743. doi: 10.1007/s00103-007-0235-5

Maayan, L., Hoogendoorn, C., Sweat, V., and Convit, A. (2011). Disinhibited eating in obese adolescents is associated with orbitofrontal volume reductions and executive dysfunction. *Obesity* 19, 1382–1387. doi: 10.1038/oby.2011.15

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Mond, J. M., Stich, H., Hay, P. J., Kraemer, A., and Baune, B. T. (2007). Associations between obesity and developmental functioning in pre-school children: a population-based study. *Int. J. Obes.* 31, 1068–1073. doi: 10.1038/sj.ijo.0803644

Monterosso, J., Ehrman, R., Napier, K. L., O'Brien, C. P., and Childress, A. R. (2001). Three decision-making tasks in cocaine-dependent patients: do they measure the same construct? *Addiction* 96, 1825–1837. doi: 10.1046/j.1360-0443.2001.9612182512.x

Moß, A., Wabitsch, M., Kromeyer-Hauschild, K., Reinehr, T., and Kurth, B. M. (2007). Prävalenz von Übergewicht und Adipositas bei deutschen Einschulkindern [Prevalence of overweight and obesity in German first graders]. *Bundesgesundheitsblatt* 50, 1424–1431. doi: 10.1007/s00103-007-0370-z

Muthén, L. K., and Muthén, B. O. (1998–2012). *Mplus User's Guide, 7th Edn*. Los Angeles, CA: Muthén and Muthén.

Nederkoorn, C., Braet, C., Van Eijs, Y., Tanghe, A., and Jansen, A. (2006). Why obese children cannot resist food: the role of impulsivity. *Eat. Behav.* 7, 315–322. doi: 10.1016/j.eatbeh.2005.11.005

Ogden, C. L., Carroll, M. D., Curtin, L. R., McDowell, M. A., Tabak, C. J., and Flegal, K. M. (2006). Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA* 295, 1549–1555. doi: 10.1001/jama.295.13.1549

Overman, W. H., Bachevalier, J., Schuhmann, E., and Ryan, P. (1996). Cognitive gender differences in very young children parallel biologically based cognitive gender differences in monkeys. *Behav. Neurosci.* 110, 673–684. doi: 10.1037/0735-7044.110.4.673

Passler, M. A., Isaac, W., and Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Dev. Neuropsychol.* 1, 349–370. doi: 10.1080/87565648509540320

Petermann, F., and Petermann, U. (eds.). (2007). *Hamburg-Wechsler-Intelligenztest für Kinder IV (HAWIK-IV) [Wechsler Intelligence Scale for Children IV]*. Göttingen: Hogrefe

Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., and Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *J. Exp. Child Psychol.* 108, 621–637. doi: 10.1016/j.jecp.2010.09.008

Reavis, R., and Overman, W. H. (2001). Adult sex differences on a decision-making task previously shown to depend on the orbital prefrontal cortex. *Behav. Neurosci.* 115, 196–206. doi: 10.1037/0735-7044.115.1.196

Reynolds, B., and Schiffbauer, R. (2005). Delay of gratification and delay discounting: a unifying feedback model of delay-related impulsive behavior. *Psychol. Rec.* 55, 439–460.

Roebers, C. M., Röthlisberger, M., Cimeli, P., Michel, E., and Neuenschwander, R. (2011). School enrolment and executive functioning: A longitudinal perspective on developmental changes, the influence of learning context, and the prediction of pre-academic skills. *Eur. J. Dev. Psychol.* 8, 526–540. doi: 10.1080/17405629.2011.571841

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain Cogn.* 55, 11–29. doi: 10.1016/S0278-2626(03)00277-X

Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol. Res. Online* 8, 23–74.

Sleddens, E. F., Kremers, S. P., and Thijs, C. (2008). The children's eating behaviour questionnaire: factorial validity and association with Body mass index in Dutch children aged 6–7. *Int. J. Behav. Nutr. Phys. Act.* 5, 1–9. doi: 10.1186/1479-5868-5-49

Smith, E., Hay, P., Campbell, L., and Trollor, J. N. (2011). A review of the association between obesity and cognitive function across the lifespan: implications for novel approaches to prevention and treatment. *Obes. Rev.* 12, 740–755. doi: 10.1111/j.1467-789X.2011.00920.x

Smith-Donald, R., Raver, C. C., Hayes, T., and Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Child. Res. Q.* 2, 173–187. doi: 10.1016/j.ecresq.2007.01.002

Sonuga-Barke, E. J. S., Dalen, L., and Remington, B. (2003). Do executive deficits and delay aversion make independent contributions to preschool attention-deficit/hyperactivity disorder symptoms? *J. Am. Acad. Child Adolesc. Psychiatry* 42, 1335–1342. doi: 10.1097/01.chi.0000087564.34977.21

Spinella, M., and Lyke, J. (2004). Executive personality traits and eating behavior. *Int. J. Neurosci.* 114, 83–93. doi: 10.1080/00207450490249356

Stadler, C., Janke, W., and Schmeck, K. (2004). *Inventar zur Erfassung von Impulsivität, Risikoverhalten und Empathie bei 9-14 Jährigen Kindern [Inventory for the Assessment of Impulsivity, Risk Behavior and Empathy in Children Aged 9-14 Years]*. Göttingen: Hogrefe.

Stevens, J. P. (2001). *Applied Multivariate Statistics for the Social Sciences, 4th Edn.* Mahwah, NJ: Lawrence Erlbaum Associates

Tataranni, P. A., Gautier, J.-F., Chen, K., Uecker, A., Bandy, D., Salbe, A. D., et al. (1999). Neuroanatomical correlates of hunger and satiation in humans using positron emission tomography. *Proc. Natl. Acad. Sci. U.S.A.* 96, 4569–4574. doi: 10.1073/pnas.96.8.4569

Tchanturia, K., Davies, H., and Campbell, I. C. (2007). Cognitive remediation therapy for patients with anorexia nervosa: preliminary findings. *Ann. Gen. Psychiatry* 6, 14. doi: 10.1186/1744-859X-6-14

Van der Merwe, M.-T. (2007). Psychological correlates of obesity in women. *Int. J. Obes.* 31, 14–18. doi: 10.1038/sj.ijo.0803731

Van Strien, T., Frijters, J. E. R., Bergers, G. P. A., and Defares, P. B. (1986). The Dutch eating behavior questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *Int. J. Eat. Disord.* 5, 295–315.

Van Strien, T., and Oosterveld, P. (2008). The children's DEBQ for assessment of restrained, emotional, and external eating in 7−to 12−year−old children. *Int. J. Eat. Disord.* 41, 72–81. doi: 10.1002/eat.20424

Volkow, N. D., Wang, G.-J., Telang, F., Fowler, J. S., Goldstein, R. Z., Alia-Klein, N., et al. (2009). Inverse association between BMI and prefrontal metabolic activity in healthy adults. *Obesity* 17, 60–65. doi: 10.1038/oby.2008.469

Wardle, J., Guthrie, C. A., Sanderson, S., and Rapoport, L. (2001). Development of the children's eating behaviour questionnaire. *J. Child Psychol. Psychiatry* 42, 963–970. doi: 10.1111/1469-7610.00792

Waxman, S. E. (2009). A systematic review of impulsivity in eating disorders. *Eur. Eat. Disord. Rev.* 17, 408–425. doi: 10.1002/erv.952

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., and Espy, K. A. (2011). The structure of executive function in 3-year-olds. *J. Exp. Child Psychol.* 108, 436–452. doi: 10.1016/j.jecp.2010.08.008

Willoughby, M., Kupersmidt, J., Voegler-Lee, M., and Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Dev. Neuropsychol.* 36, 162–180. doi: 10.1080/87565641.2010.549980

World Health Organization Europe. (2006). *European Charter on Counteracting Obesity. WHO European Ministerial Conference on Counteracting Obesity*. Istanbul. Available online at: http://www.euro.who.int/en/health-topics/noncommunicable-diseases/obesity/publications/pre-2009/european-charter-on-counteracting-obesity

Wulfert, E., Block, J. A., Ana, E. S., Rodriquez, M. L., and Colsman, M. (2002). Delay of gratification: impulsive choices and problem behaviors in early and late adolescence. *J. Pers.* 70, 533–552. doi: 10.1111/1467-6494.05013

Zelazo, P. D., and Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: development and plasticity. *Child Dev. Perspect.* 6, 354–360. doi: 10.1111/j.1750-8606.2012.00246.x

Zelazo, P. D., Carlson, S. M., and Kesek, A. (2008). "Development of executive function in childhood," in *Handbook of Developmental Cognitive Neuroscience, 2nd Edn.*, eds C. A. Nelson and M. Luciana (Cambridge, MA: MIT Press), 553–574.

Zelazo, P. D., Carter, A., Reznick, J. S., and Frye, D. (1997). Early development of executive function: a problem-solving framework. *Rev. Gen. Psychol.* 1, 198–226. doi: 10.1037/1089-2680.1.2.198

Zelazo, P. D., and Müller, U. (2002). "Executive function in typical and atypical development," in *Handbook of Childhood Cognitive Development*, ed U. Goswami (Oxford: Blackwell), 445–469.

Zhang, L., Karabenick, S. A., Maruno, S., and Lauermann, F. (2011). Academic delay of gratification and children's study time allocation as a function of proximity to consequential academic goals. *Learn. Instr.* 21, 77–94. doi: 10.1016/j.learninstruc.2009.11.003

Zimmermann, P., Gondan, M., and Fimm, B. (2002). *KITAP Testbatterie zur Aufmerksamkeitsprüfung für Kinder [KITAP Attention Test Battery for Children]*. Herzogenrath: Psychologische Testsysteme.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2014; accepted: 27 April 2014; published online: 19 May 2014.
*Citation: Groppe K and Elsner B (2014) Executive function and food approach behavior in middle childhood. Front. Psychol. **5**:447. doi: 10.3389/fpsyg.2014.00447*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Groppe and Elsner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 07 April 2014
doi: 10.3389/fpsyg.2014.00285

Getting the right grasp on executive function

Claudia L. R. Gonzalez[1]*, Kelly J. Mills[1], Inge Genee[2], Fangfang Li[3], Noella Piquette[4], Nicole Rosen[2] and Robbin Gibb[5]

[1] Department of Kinesiology, The Brain in Action Laboratory, University of Lethbridge, Lethbridge, AB, Canada
[2] Department of Modern Languages, University of Lethbridge, Lethbridge, AB, Canada
[3] Department of Psychology, Univeristy of Lethbridge, Lethbridge, AB, Canada
[4] Department of Education, University of Lethbridge, Lethbridge, AB, Canada
[5] Department of Neuroscience, University of Lethbridge, Lethbridge, AB, Canada

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Vanessa R. Simmering, University of Wisconsin Madison, USA
Marianne Jover, Aix marseille University, France

***Correspondence:**
Claudia L. R. Gonzalez, Department of Kinesiology, The Brain in Action Laboratory, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada
e-mail: claudia.gonzalez@uleth.ca

Executive Function (EF) refers to important socio-emotional and cognitive skills that are known to be highly correlated with both academic and life success. EF is a blanket term that is considered to include self-regulation, working memory, and planning. Recent studies have shown a relationship between EF and motor control. The emergence of motor control coincides with that of EF, hence understanding the relationship between these two domains could have significant implications for early detection and remediation of later EF deficits. The purpose of the current study was to investigate this relationship in young children. This study incorporated the Behavioral Rating Inventory of Executive Function (BRIEF) and two motor assessments with a focus on precision grasping to test this hypothesis. The BRIEF is comprised of two indices of EF: (1) the Behavioral Regulation Index (BRI) containing three subscales: Inhibit, Shift, and Emotional Control; (2) the Metacognition Index (MI) containing five subscales: Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. A global executive composite (GEC) is derived from the two indices. In this study, right-handed children aged 5–6 and 9–10 were asked to: grasp-to-construct (Lego® models); and grasp-to-place (wooden blocks), while their parents completed the BRIEF questionnaire. Analysis of results indicated significant correlations between the strength of right hand preference for grasping and numerous elements of the BRIEF including the BRI, MI, and GEC. Specifically, the more the right hand was used for grasping the better the EF ratings. In addition, patterns of space-use correlated with the GEC in several subscales of the BRIEF. Finally and remarkably, the results also showed a reciprocal relationship between hand and space use for grasping and EF. These findings are discussed with respect to: (1) the developmental overlap of motor and executive functions; (2) detection of EF deficits through tasks that measure lateralization of hand and space use; and (3) the possibility of using motor interventions to remediate EF deficits.

Keywords: grasping movements, left hemisphere, space use, development, frontal lobe, handedness, assessment, intervention

INTRODUCTION

Historically, neuropsychological evidence has highlighted the role of the frontal cortex in the planning and execution of behavior (Kolb and Whishaw, 2009). Patients with frontal lobe injury present with a host of motor and cognitive disturbances. In the motor domain, frontal lobe injury could lead to deficits in gross motor function (e.g., impaired posture and gait) and/or fine motor control (e.g., impaired reaching and grasping). In the cognitive domain some of the most commonly disrupted functions include: initiation, planning, purposive action, self-monitoring, self-regulation, and volition (Stuss, 2011). This has led to the understanding that the frontal lobe is the area that supports executive function (EF). EF is a blanket term that is considered to include attentional control, self-regulation, inhibition, working memory, goal setting, planning, problem solving, mental flexibility, and abstract reasoning (Diamond and Lee, 2011).

Early in life, children learn and refine a host of motor skills that will have a phenomenal impact on later cognitive function. In fact, there is evidence that the time scales for development of these functions imbricate (see Diamond, 2000; for a review, Diamond, 2007). In addition, imaging studies have shown overlapping activation of motor function and EF in the frontal lobe, in particular the dorsal premotor cortex, which responds to planning, selection, organization, and execution of actions (Abe and Hanakawa, 2009; Hanakawa, 2011). In a retrospective study Piek et al. (2008) correlated data gathered in the preschool years using the Ages and Stages Questionnaire (ASQ) for gross motor trajectory with later performance on the Wechsler Intelligence Scale in elementary school. They found a high correlation between the two, once socioeconomic status was controlled for. Furthermore, they showed a predictive relationship between motor outcomes and working memory function. They and others have concluded

that abnormalities in motor performance may be an important basis for the detection of later cognitive impairments (Piek et al., 2008; Butcher et al., 2009; Iverson, 2010). In fact, Kirby et al. (2008) report that more than 50% of university and college students with motor difficulties also suffer from difficulties with executive function. This evidence highlights the enduring nature of the relationship between motor and executive function.

An emerging research field is providing evidence of the interrelatedness of motor and executive functions, particularly in the planning domain (Pennequin et al., 2010; Thibaut and Toussaint, 2010; van Swieten et al., 2010; Jongbloed-Pereboom et al., 2013; and see Rosenbaum et al., 2012 for a review). For example, recently Jongbloed-Pereboom et al. (2013) asked 3–10 years old children to grasp a wooden sword and place it into a fitted aperture. The handle of the sword was placed in one of six different orientations. The authors documented the grip type that participants used and analyzed it with respect to end-state comfort. It was found that action planning increased from 3 to 10 years of age. Ten year olds behaved more like adults such that they preferred an awkward initial grasp to assure a final end-state comfort. Authors conclude that a cognitive component directly related to anticipatory planning subserves the performance of this task. Given that both planning and inhibition are critical components of EF, this evidence suggests a rich connection between cognition and action. Based on this literature, we hypothesized that measures of motor performance and EF could be mutually predictive. A motor action that we perform hundreds of times each day is reaching and grasping. Grasping has been shown to develop as early as 6 months of age and can be reliably assessed by age one (Michel et al., 2006; Jacquet et al., 2012; Sacrey et al., 2012, 2013). Using such an ecologically-valid measure of motor performance we sought to investigate its possible relationship with EF. If this relationship is established, the implications are paramount for improving life-long success, for three reasons. First, skilled motor ability can be readily assessed earlier than EF. Second, EF has been shown to be a better predictor of school success than IQ (Blair and Razza, 2007; Diamond and Lee, 2011; Masten et al., 2012). Third, if developmental delays are detected, interventions for both motor skill and EF training can be implemented immediately to prevent academic setbacks later in life.

In the present investigation we examined EF and motor performance in two groups of children; 5–6 and 9–10 year olds. We used the Behavioral Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000) to assess EF and two reaching and grasping tasks to assess motor performance. The BRIEF was developed as an ecologically valid model to assess children's executive functions (Gioia et al., 2000). According to Gioia and Isquith (2004), the BRIEF was designed as "a means of culling and standardizing the rich information provided by parents and teachers in a more reliable and efficient manner with known psychometric properties." This test has been widely used to assess executive function in normal and clinical populations and there have been several validity studies demonstrating its effectiveness (for review see Donders, 2002; Strauss, 2006). Moreover, a recent study corroborated the effectiveness of the BRIEF as a tool to assess EF, as it was found that BRIEF measures correlated with in-lab behavioral measures (Lalonde et al., 2013). Furthermore, studies have shown strong correlations with academic performance and scores obtained with the BRIEF (e.g., Waber et al., 2006).

Reaching and grasping was assessed using two well-studied grasping tasks: grasp-to-place and grasp-to-construct (Gonzalez et al., 2006, 2007; Gonzalez and Goodale, 2009; Gallivan et al., 2011; Sacrey et al., 2013; Stone et al., 2013; Stone and Gonzalez, 2014). In the grasp-to-place task participants are asked to reach for and grasp wooden blocks with colors or numbers and place them into a box. The grasp-to-construct task requires individuals to locate, reach for and grasp plastic blocks (LEGO®) of different size, shape, and color in order to replicate a model based on a sample. Because the grasp-to-construct task demands that participants plan and strategize in order to reproduce the sample as fast and accurately as possible, we hypothesized that this task, in particular, would be sensitive to a relationship between motor and executive function.

MATERIALS AND METHODS

PARTICIPANTS

A total of 40 children took part in the study. All children were identified as right-handed according to a modified version of the Edinburgh Handedness questionnaire (Oldfield, 1971; completed by each parent; see Stone et al., 2013 for full version of the questionnaire). Thirty-one children had previously participated in a psychological study at the University of Lethbridge (U of L), at which time their parents had opted to receive e-mail notifications of future studies at the U of L. The remaining children were recruited through either acquaintances of the authors, or at a booth during a public children's festival. Nineteen individuals comprised the "younger" age group of 5 and 6 year olds (11 females; M $\pm$ *SD* age $=$ 5.98 $\pm$ 0.53 years) and 21 individuals comprised the "older" age group of 9 and 10 year olds (10 females; M $\pm$ *SD* age $=$ 9.88 $\pm$ 0.51 years). Participants were healthy, with no evidence of neurological impairment. Participants were naïve to the purpose of the study and informed parental consent, as well as child verbal consent, was obtained prior to participation.

PROCEDURE

Parent questionnaires

After informed consent was obtained, the parent accompanying the child participant was given three paper-based questionnaires to be completed: (1) a participant information sheet that consisted of general questions regarding the child's motor, cognitive, and language development. (2) a modified version of the Edinburgh Handedness Inventory (to be filled out with the child's hand preferences in mind); and (3) the BRIEF (Gioia et al., 2000). For the BRIEF, the parent was asked to rate 86 everyday behaviors over the past 6 months as never occurring, sometimes occurring, or often a problem for their child. Each behavior belongs to one of eight subscales that represent unique facets of executive function (Gioia and Isquith, 2004): (1) Inhibit (resist or delay an impulse); (2) Shift (change problem-solving strategies); (3) Emotional Control (appropriately modulate affective reactivity); (4) Initiate (begin a task or activity, generate ideas); (5) Working Memory (hold information in mind for the purpose of completing a task); (6) Plan/Organize (anticipate events, set goals, and

develop steps to carry out a task); (7) Organization of Materials (establishing and maintaining order to systematically carry out a task); (8) Monitor (check action to assure appropriate attainment of a goal). Scores for each subscale were obtained by summing the parent's score of each item for each subscale. The first three subscales were summed to comprise the Behavioral Regulation Index (BRI), while the next five were summed to comprise the Metacognitive Index (MI). Together the two indices form the Global Executive Composite (GEC; the child's overarching score of executive function). The BRIEF includes built-in checks for parent negativity and inconsistency in responses. The raw scores obtained from the eight subscales, two indices, and GEC are converted to standard scores based on age and gender norms provided in the BRIEF handbook (Gioia et al., 2000). In the present study, both raw and standardized scores were subjected to statistical analysis.

While the parent completed the three questionnaires in an area outside the testing lab, the child was welcomed into the lab with a "treasure map" and told that he/she could find a treasure by playing a few games (motor tasks) with the experimenter. The child participated in two tasks: grasp-to-construct and grasp-to-place. The tasks occurred in the same order for all participants. Tasks were video recorded with a JVC Everio HD camera positioned directly in front of the work-space, facing the seated participant and aligned with his/her midline. All children sat in chairs without armrests, and no directions were ever given regarding hand use.

Grasp-to-construct

The child was asked to sit and face a table, with a workspace covered in Lego® blocks. The workspace was notionally divided into four quadrants of equal dimensions: left near (LN), left far (LF), right near (RN), and right far (RF). Each of the 4 quadrants contained the exact same set of pieces, which were unique in size, shape, and color within the set (see **Figure 1A**). In this task, the child was required to replicate four pre-made models. Each one was comprised of one set of pieces (the same set placed in each quadrant); thus, models contained the same pieces but in unique configurations. Within each age group, all children received the same four models, in the same order. The four sets of pieces on the table were placed in near-mirror image positions relative to one another, so that there was an equal opportunity to choose pieces from LN, LF, RN, or RF space when completing the models.

Individuals in the younger group (5–6 years old) sat at a table with a workspace 60 cm deep × 80 cm wide. These children encountered a total of 20 pieces on the tabletop; each of the four quadrants and four models contained the same set of five pieces. The older group (9–10 years old) sat at a table with a workspace 70 cm deep × 122 cm wide. These children encountered a total of 40 pieces (each quadrant and model contained the same set of 10 pieces).

Once seated, the experimenter explained to the child that the object of the "game" was to make a model that looked just like the experimenter's model. The experimenter gestured to a pre-made model, placed across from the child at the far end of the block array, aligned with the child's midline (see **Figure 1A**). Children in the older age group only were asked to complete the replica as quickly as possible. Children were allowed to pick up the original model at any point during the task, and manipulate it in any way to understand its configuration. However, models were designed to be fully understood from a straight-on viewing angle (see **Figure 1B** for an example). Once the first replica was complete, the experimenter removed the replica and replaced the first model with the next (in the same position). At the onset of the second trial, three sets of pieces were still available on the tabletop. After completion of all four replicas, all pieces on the table-top were used.

Grasp-to-place

Immediately after the completion of the grasp-to-construct task, the child was seated at a table on which a total of 40 numbered and 20 colored blocks (2.54 cm^3) were arranged in a rectangular array of six rows and 10 columns (see **Figure 2**). Blocks were placed approximately 6.35 cm apart, creating a grid approximately 33 cm deep × 61 cm wide. The grid was notionally divided into right and left space. One set of blocks (presented on one half of space) contained 20 blocks labeled with the numbers 0–19 and 10 blocks of different colors; blocks were placed in pseudo-random positions. In the other half of space, a replicate set of blocks was placed in a near-symmetric fashion. The placement of all 60 blocks was consistent across participants. At the far end of the array, a cardboard box 31.5 cm wide by 21.5 cm deep and decorated to look like a "monster's mouth" was placed.

The experimenter told the child that she was going to read a list of numbers and colors out loud. After each number or color, the child was to find and pick up one and only one corresponding block, and place it into the box. All participants were encouraged to be as fast as possible and no instruction as to what hand/space to use was given. Each number (0–19) and eight colors (28 requests total) were called out once in a pseudo-random order.

DATA PROCESSING AND ANALYSIS

Brief

The BRIEF was scored according to scoring procedures outlined in the BRIEF handbook (Gioia et al., 2000). For each child, raw and standard scores were obtained for each component: the GEC, two indices (BRI and MI), and eight subscales.

Grasping tasks

All video recordings were analyzed offline.

Time-to-complete

Total latency to complete the four models in the grasp-to-construct task and the time required to place the numbered and colored bocks in the grasp-to-place task was recorded.

Hand use

Within each task, the hand used (left or right) for every grasp to a target item—a Lego® block or wooden block—was scored. The total number of grasps was calculated to determine the percentage of right hand use [(number of grasps with right hand/total number of grasps) ×100] for each individual on each task.

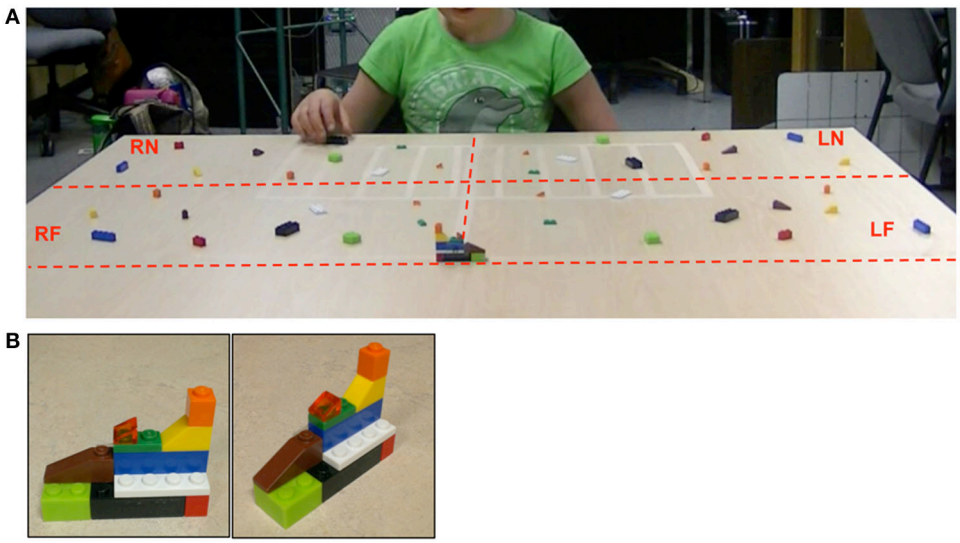


FIGURE 1 | (A) The picture illustrates the workspace used by older children (9 and 10 years old) in the grasp-to-construct task. The table was notionally divided into four quadrants of equal dimensions (lines were not visible). Four identical sets of 10 pieces were placed on the tabletop—one set in each quadrant in near-mirror image placements. Within a set, pieces were unique in color and shape. The model to be replicated on each trial was placed at the far border of the workspace, aligned with the child's midline. **(B)** The figure demonstrates the first model older children were prompted to replicate in the grasp-to-construct task, from straight-on and side view angles. Each of the four models was composed of one piece set (contained in each quadrant on the table). Models were arranged such that they could be fully understood from a straight-on viewing angle, however, participants were allowed to pick up and rotate the model at any point during construction.

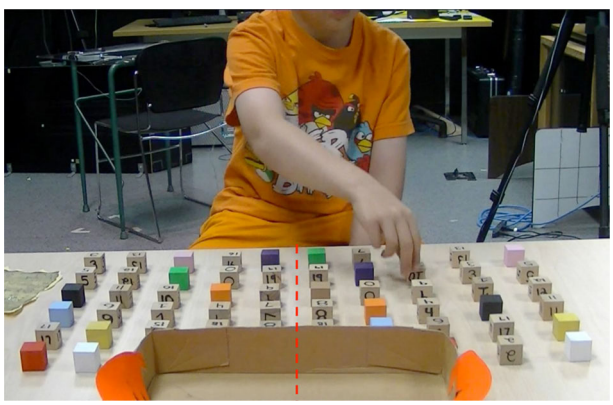


FIGURE 2 | An illustration of the workspace in the grasp-to-place task. The table was notionally divided into left and right space; 2 identical sets of 20 numbered and 10 colored blocks were placed in left and right space in near-mirror image positions that remained consistent across participants. The experimenter called out a pseudo-random list of numbers and colors; after each, the child was to locate one correspondingly-labeled block as quickly as possible, and place it into the box at the far end of the array (the "monster's mouth").

Space use in the grasp-to-construct task

In a previous study from Gonzalez' lab using the grasp-to-construct task with adults (de Bruin et al., 2014), differential use of space for grasping (left vs. right and near vs. far) was shown. Right-handed participants grasp from right-near space earlier than anywhere else. We explored the possibility that adult-typical patterns of space use in children would be correlated with better EF. Space use in the grasp-to-construct task was investigated by assigning a number to each grasp based on the order in which the grasp occurred (the first grasp received a 1, the second a 2, the third a 3, and so forth). At task completion, each quadrant had five grasp values assigned to it for the younger group. For example, if the first five grasps made by a participant occurred in the right near quadrant, the values 1, 2, 3, 4, and 5 would be assigned to that quadrant. Within each quadrant, values were then summed to produce four quadrant sums and two hemi-space sums (L and R). The lowest possible quadrant sum for the younger group was 15 (1, 2, 3, 4, and 5), and the highest possible sum was 90 (16, 17, 18, 19, 20). In the older group, 10 pieces were placed in each quadrant, raising the minimum quadrant sum to 55 and the maximum to 355. Each quadrant and hemi-space sum was then divided by the table sum (210 in the younger group, 820 in the older group), to obtain quadrant and hemi-space percentages. The lower the percentage for a given space, the earlier in the task that space was attended to and exhausted of pieces (de Bruin et al., 2014).

All data were analyzed using SPSS Statistics 19.0 for Mac (SPSS Inc., Chicago, IL, USA). Statistical significance was set at $\alpha = 0.05$. Correlation (Pearson's r) and regression analyses (linear) between scores from the BRIEF and scores from the grasping tasks were computed. In addition, means and standard errors for the time-to-complete and hand use for grasping are reported below. The results were analyzed for overall effects (both age groups together) and then inspected separately for each age group. Only significant results are reported.

RESULTS

No statistically significant differences were found with respect to sex in either age group or in any of the measurements, therefore the data was collapsed across this variable.

DESCRIPTIVE STATISTICS

In the BRIEF, lower scores are associated with better EF. **Table 1** shows the results for children in the two age groups for each of the components of the BRIEF. In the grasp-to-construct task the younger group spent on average 141.42 ± 10.41 (SEM) s completing the task whereas the older group spent on average 191.95 ± 8.4 s. The older group required more time to complete the task because they were presented with 40 Lego® blocks instead of the 20 blocks the younger group worked with. In the grasp-to-place task the younger group spent on average 250.73 ± 15.41 s completing the task whereas the older group spent on average 114.95 ± 5.0 s. In this case, both groups were presented with the same number of wooden blocks.

In the grasping tasks, both groups of children displayed a right hand preference. In the grasp-to-construct task, percent right hand use in the younger children was 59.82 ± 12.42 and the older children 68.11 ± 14.23. In the grasp-to-place task these values were 74.47 ± 23.74 and 85.54 ± 17.84, respectively.

Children of both ages displayed a preference for attending first to right space and specifically to right-near space. In the younger group, percent right hemispace use was 44.89 ± 10.28 and the older group was 40.44 ± 8.29. For the right near quadrant, sum averages were 15.74 ± 5.95 and 17.65 ± 6.25, respectively.

CORRELATION ANALYSES USING BRIEF STANDARD SCORES

Our main hypothesis was that measures of motor performance would correlate with executive function. The dependent variables in both grasping tasks were the time that participants took to complete each task and the hand used to pick up the blocks. We hypothesized that faster times in completing the tasks, particularly in the grasp-to-construct, would correlate with better EF. We had no particular prediction regarding hand use and its possible relationship with EF. In addition, space use was documented during the grasp-to-construct task to explore the possibility that children exhibiting adult-typical space use (right-handed participants attend to right-near space first) in the grasp-to-construct task would have better EF scores.

No significant correlations were found for the standard scores of the BRIEF and the time to complete either grasping task.

As mentioned previously, lower scores on the BRIEF indicate better EF. Therefore, a negative correlation between right hand use and EF would indicate that the more the right hand is used for grasping the better the EF score.

Overall (age groups combined), there was a significant negative correlation between **hand use** in the **grasp-to-construct task** and the standard score on the Inhibit subscale of the BRIEF [$r_{(40)} = -0.39$; $p < 0.02$]. A closer look at this correlation revealed the significant effect was mostly driven by the younger children [$r_{(19)} = -0.52$; $p < 0.03$]. In addition, when looking at this young group a significant correlation was also found between right hand use and the score on the Monitor subscale [$r_{(19)} = -0.62$; $p < 0.01$]. Furthermore, trends were noted for Emotional Control [$r_{(19)} = -0.41$, $p = 0.09$], BRI [$r_{(19)} = -0.45$, $p = 0.05$], and GEC [$r_{(19)} = -0.41$, $p = 0.08$]. No other significant correlations were found for any of the remaining subscales or age groups.

For the **grasp-to-place task** overall, there was a significant negative correlation between **hand use** and the standard GEC score [$r_{(40)} = -0.37$; $p < 0.02$; see **Figure 3A**]. Furthermore, the correlation was maintained across the two indices; BRI [$r_{(40)} = -0.33$; $p < 0.05$] and MI [$r_{(40)} = -0.35$; $p < 0.05$]. Closer examination revealed significant correlations for Inhibit [$r_{(40)} = -0.43$; $p < 0.01$], Working Memory [$r_{(40)} = -0.32$; $p < 0.05$], Plan [$r_{(40)} = -0.35$; $p < 0.05$], and Monitor [$r_{(40)} = -0.42$; $p < 0.01$]. When separated by age, the correlation held for Monitor [$r_{(19)} = -0.54$; $p < 0.02$] and a trend for Plan was observed [$r_{(19)} = -0.40$; $p = 0.09$] in the younger group. For the older group, a trend was observed for Inhibit [$r_{(21)} = -0.40$; $p = 0.07$].

As previously stated, we explored the possibility that children exhibiting adult-typical **space use** in the grasp-to-construct task would have better EF scores. Lower scores on any space sum (%) are indicative of children attending to that space earlier (see Materials and Methods). A positive correlation between space sum and the scores of the BRIEF indicate that the earlier a child attends to that space, the better the EF. Results showed that the earlier the **right hemispace** was attended to, the better the EF score. Overall there was a significant positive correlation between **right hemispace sum (%)** and the standard GEC score [$r_{(40)} = 0.33$; $p < 0.05$] (see **Figure 3B**). Closer examination revealed a significant positive correlation for Plan [$r_{(40)} = 0.36$; $p < 0.05$]. In addition, trends were observed for the two indices; BRI [$r_{(40)} = 0.31$; $p = 0.057$] and MI [$r_{(40)} = 0.30$; $p = 0.068$], and the subscales Inhibit [$r_{(40)} = 0.31$; $p = 0.058$], Working Memory [$r_{(40)} = 0.28$; $p = 0.083$], and Monitor [$r_{(40)} = 0.31$; $p = 0.054$]. These effects were mostly driven by the older group. For this group there was a significant positive correlation between **right hemispace sum (%)** and the standard GEC score [$r_{(21)} = 0.59$;

Table 1 | Mean standard scores and standard deviations on the eight subscales, two indices, and General Executive Composite of the BRIEF, for all participants and the two separate age groups.

BRIEF component	All ages	Younger	Older
General Executive Composite (GEC)	53.65 (±10.56)	56.05 (±10.74)	51.48 (±10.16)
Behavior Regulation Index (BRI)	54.35 (±10.4)	58.11 (±10.47)	50.95 (±9.32)
Inhibit	53.05 (±12.11)	57.68 (±14.03)	48.86 (±8.37)
Shift	54.9 (±12.21)	57.21 (±12.72)	52.81 (±11.63)
Emotional control	53.75 (±10.27)	56.63 (±8.86)	51.14 (±10.96)
Metacognitive Index (MI)	52.7 (±10.69)	53.58 (±10.53)	51.9 (±11.03)
Initiate	52.35 (±10.13)	52.26 (±10.94)	52.43 (±9.6)
Working memory	54.55 (±11.07)	56.79 (±10.97)	52.52 (±11.03)
Plan/Organize	51.1 (±9.02)	51.56 (±7.18)	50.71 (±10.5)
Organization of materials	53.55 (±10.1)	54.79 (±9.54)	52.43 (±10.7)
Monitor	51.45 (±11.56)	53.63 (±11.96)	49.48 (±11.09)

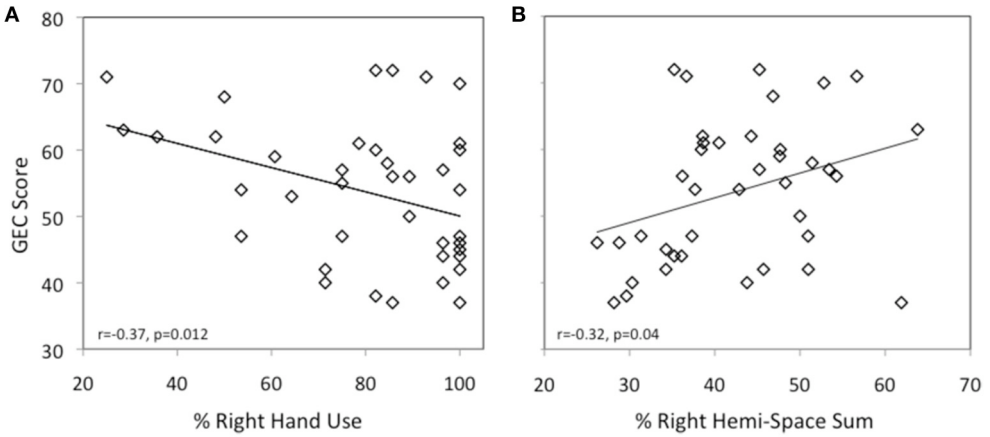


FIGURE 3 | (A) The graph depicts the relationship between percent right hand use in the grasp-to-place task and the standard score obtained on the General Executive Composite of the BRIEF for all children (younger and older). A significant negative correlation was observed ($r = -0.368$, $p = 0.019$), indicating that the more the right hand was used for grasping, the lower (better) the overarching EF score. **(B)** The graph depicts the relationship between percent right hemi-space sum in the grasp-to-place task and the standard score obtained on the General Executive Composite of the BRIEF for all children (younger and older). A smaller percent sum indicates earlier attendance to the right space. A significant positive correlation was observed ($r = 0.327$, $p = 0.042$), demonstrating that the earlier the right space was attended to, the lower (better) the overarching EF score.

$p < 0.01$]. Significant positive correlations were also found for MI [$r_{(21)} = 0.61$; $p < 0.005$], and the subscales Initiate [$r_{(21)} = 0.43$; $p = 0.05$], Working Memory [$r_{(21)} = 0.61$; $p < 0.005$], Plan [$r_{(21)} = 0.65$; $p < 0.005$], and Organization of Materials [$r_{(21)} = 0.44$; $p < 0.05$]. Trends were observed for BRI [$r_{(21)} = 0.39$; $p = 0.08$], and the subscales Inhibit [$r_{(21)} = 0.39$; $p = 0.079$], and Monitor [$r_{21)} = 0.43$; $p = 0.05$]. Again, the earlier the right space was attended to, the better the EF score. We further investigated the hemi-space effect in the older group by looking at the **right near quadrant space use (%)** and found that the earlier the right near quadrant was attended to, the better the EF score. Consistent with our hypothesis, significant positive correlations between right-near space sum (%) were found for the standard GEC score [$r_{(21)} = 0.57$; $p < 0.01$], MI [$r_{(21)} = 0.56$; $p < 0.005$], Inhibit [$r_{(21)} = 0.46$; $p < 0.05$], Initiate [$r_{(21)} = 0.51$; $p < 0.02$], Working Memory [$r_{(21)} = 0.63$; $p < 0.005$], Plan [$r_{(21)} = 0.57$; $p < 0.01$], and Organization of Materials [$r_{(21)} = 0.53$; $p < 0.02$].

CORRELATION ANALYSES USING BRIEF RAW SCORES

Because it is known that EF improves with developmental age (for a review see Best and Miller, 2010) we wondered whether right hand use increases as well-with developmental age and if therefore our results could be explained on the basis of age alone. In other words, we investigated whether the relationship between hand use and EF score is an epiphenomenon of hand use changing with age (i.e., whether children get more right-handed as they age). We found no significant correlation between chronological age (days) and right hand use in either grasping task: grasp-to-construct [$r_{(40)} = 0.39$; $p > 0.05$] or grasp-to-place [$r_{(40)} = 0.24$; $p > 0.1$].

Given that the BRIEF standardizes raw scores to normative data for age, we explored possible correlations between chronological age (days) and raw BRIEF scores. We found a significant negative correlation between chronological age and the BRI [$r_{(40)} = -0.34$; $p < 0.05$] as well as the Inhibit [$r_{(40)} = -0.41$; $p < 0.01$] subscale of the BRIEF; the older the child the better their EF score.

Unexpectedly, we found more significant correlations between the BRIEF raw scores and hand use, than the BRIEF raw scores and age. Overall (both ages combined), there was a significant correlation between right hand use in the grasp-to-construct task and the raw scores on the Inhibit subscale [$r_{(40)} = -0.44$; $p < 0.005$]. In the grasp-to-place task, right hand use correlated raw scores on the GEC [$r_{(40)} = -0.36$; $p < 0.05$], MI [$r_{(40)} = -0.32$; $p < 0.05$], BRI [$r_{(40)} = -0.34$; $p < 0.05$], Monitor [$r_{(40)} = -0.37$; $p < 0.02$], and Inhibit [$r_{(40)} = -0.47$; $p < 0.002$]. Further analysis revealed that the observed correlations were mostly driven by the younger group. Within this group significant correlations were found for Inhibit [$r_{(19)} = -0.577$, $p = 0.01$], BRI [$r_{(19)} = -0.498$, $p = 0.03$], Monitor [$r_{(19)} = -0.614$, $p = 0.007$], borderline GEC [$r_{(19)} = -0.444$, $p = 0.057$], and Emotional Control [$r_{(19)} = -0.409, p = 0.082$].

REGRESSION ANALYSES

To explore the contributions that age, hand-use, and space-use had on EF we conducted several linear regression analyses. Given that the grasp-to-place task yielded more and stronger correlations of right-hand use with EF, we used this measure in the hand use regression analyses. For space-use, right near quadrant sum was used in the computation. For simplicity we focused on the GEC as the dependent measure. The model accounted for 15.7% of the variance, and it was significant [$F_{(3, 39)} = 3.4$; $p < 0.05$]. An examination of the coefficients showed that right hand use and right-near space use were significant predictors of EF (see **Table 2**). Interestingly, age was not a predictor of EF. To explore the possibility of a mutually predictive relationship,

Table 2 | Results of the regression analyses. Note the relationship between hand and space use during the grasping tasks and EF.

Dependent measure		Coefficients							
		Unstandardized coefficients		Standardized coefficients				Correlations	
		B	Std. error	Beta	*t*	Sig.	Zero-order	Partial	Part
GEC									
	Chrono-age	−0.013	0.008	−0.39	−1.61	0.11	−0.08	−0.25	−0.23
	RH-use	−41.97	17.41	−0.36	−2.40	**0.02**	−0.36	−0.37	−0.35
	RN-space use	0.180	0.088	0.49	2.05	**0.04**	0.11	0.32	0.30
RIGHT-HAND USE									
	Chrono_age	0.00	0.00	0.04	0.17	0.86	0.23	0.03	0.02
	RN-space use	0.001	0.001	0.20	0.77	0.44	0.19	0.12	0.11
	GEC	−0.003	0.001	−0.38	−2.40	**0.02**	−0.36	−0.37	−0.36
RIGHT-NEAR SPACE USE									
	Chrono_age	0.07	0.009	0.79	8.05	**0.000**	0.79	0.80	0.77
	RH-use	25.97	33.52	0.08	0.77	0.444	0.19	0.12	0.07
	GEC	0.58	0.28	0.21	2.05	**0.047**	0.11	0.32	0.19
CHRONOLOGICAL AGE									
	RN-space use	8.83	1.09	0.81	8.05	**0.000**	0.79	0.80	0.77
	RH-use	66.07	372.1	0.01	0.17	0.86	0.23	0.03	0.01
	GEC	−5.14	3.19	−0.17	−1.61	0.11	−0.08	−0.25	−0.15

The bolded values represent the significance.

we computed a second regression analysis with right-hand use as the dependent measure and chronological age, GEC, and space use as independent measures. The model accounted for 12.0% of the variance and it was significant [$F_{(3,\ 39)} = 2.8$; $p = 0.05$]. Examination of the coefficients showed that GEC was a significant predictor of right-hand use (see **Table 2**). Neither chronological age nor space use predicted right-hand use. A final regression analysis was conducted to investigate if chronological age, hand-use and GEC would be predictors of space use. The model accounted for 64.3% of the variance and significance was noted [$F_{(3,\ 39)} = 24.4$; $p < 0.0001$]. The coefficients revealed that chronological age was a powerful predictor of right-near space use (see **Table 2**). GEC was also a predictor of space use but hand use was not.

DISCUSSION

The purpose of the present study was to investigate the possible relationship between motor performance and EF. To do this we asked children of two different ages to complete two grasping tasks while their parents filled out a questionnaire detailing their child's EF. For the grasping tasks, children reached for and grasped Lego® blocks in order to construct different models, or grasped wooden blocks to place in a box. Three aspects of their performance were assessed: the time it took them to complete each task, their preference for hand use, and their preference for space use. The results showed no relationship between EF and their performance as measured by time. In other words, how quickly a child completed the tasks bore no relationship to their scores on the BRIEF. However, the results demonstrated a robust relationship between the scores on the BRIEF and the child's preference to use their right hand and the right space for grasping. Remarkably, right hand use and right space use were predictors of EF, and EF was a reliable predictor of right hand use. These unexpected findings suggest that a more lateralized brain supports enhanced EF.

Studies have shown overlapping neural networks that support motor and EF including the frontal lobe, the cerebellum, and the basal ganglia (Schmahmann and Pandya, 2008; Abe and Hanakawa, 2009; Pangelinan et al., 2011; for a review see Diamond, 2000). At the behavioral level, numerous studies have presented evidence of motor deficits accompanying cognitive deficits (e.g., Eliason, 1986; Eliasson et al., 2004; Racine et al., 2008; Fuentes et al., 2009). Children with developmental coordination disorder for example, present with a host of gross and fine motor skill deficits. Up to 50% of these children may suffer from executive dysfunction (Willcutt and Pennington, 2000; Sugden et al., 2008) that in some cases lasts into the adult life (Kirby et al., 2008). Furthermore, fine motor skills have been used as the primary indicator of the need for intervention in kindergarten children (Roth et al., 1993). In normally developing children, studies have also reported a relationship between EF and motor performance (e.g., Roebers and Kauer, 2009; Davis et al., 2011; also Piek et al., 2008). Cameron et al. (2012) tested children in several gross and fine motor tasks and discovered that children that were better at a design copy task requiring fine motor control (copy pictures of different geometrical shapes using paper and pencil) not only performed better on tests of EF, but they also attained higher kindergarten achievement. Recently, Carlson reported that children starting kindergarten with better fine motor skill showed enhanced learning in both math and reading (Carlson et al., 2013). Based on these previous examples we hypothesized that performance measures such as time to complete the grasping task might predict EF. This was not the case. In reviewing the video footage it was obvious that individual

differences contributed to noise in this measure. For example, some children were more familiar with assembling Lego, some were very verbally interactive with the experimenter, and yet others seemed shy or introverted. These factors likely undermined the effectiveness of time as a measure of performance.

Although time to complete the grasping tasks did not correlate with any measures of the BRIEF, we found that the strength of right hand and space preference was intimately related with EF. Results from the present study suggest two potential and non-mutually exclusive scenarios: (1) the possibility that EF enjoys privileged support from the left hemisphere; and/or (2) that the greater the lateralization of function (either to the left or right hemisphere), the better the behavioral output. With respect to the first scenario, there is reasonable, albeit not explicit, evidence of increased involvement of the left hemisphere in EF. In a recent study, a large sample of brain-injured adults was subjected to neuropsychological testing and brain imaging analysis (Barbey et al., 2012). Both hemispheres were scanned for evidence of injury. Interestingly, the results showed that high-level cognitive performance (intelligence and EF) was compromised in patients with left hemisphere damage only. In a similar study of brain-damaged patients, measures of general intelligence (some of which overlap with EF) were correlated with a left lateralized fronto-parietal network (Glascher et al., 2010). Furthermore, this study identified a sector in the left anterior frontal lobe (BA 10) that was uniquely related to general intelligence. Curiously, BA 10 has also been implicated in the planning of movement (Momennejad and Haynes, 2013) and specifically a relationship has been found between better motor imagery and activation of the left "prefrontal executive" area BA10 (van der Meulen et al., 2012). In light of this evidence, it is perhaps not surprising that our participants that showed more left hemisphere lateralized biases for hand and space use also showed higher EF scores. In other words, our results provide strong evidence of left hemisphere specialization for EF.

The second possibility is that a greater degree of functional lateralization supports better motor and cognitive performance. Indeed, there is evidence to support this notion. In a study by Crow et al. (1998) 12,770 children were assessed for hand skill and cognitive control. For the hand skill task, children were given 1 min to put a check mark in as many squares as possible on a printed sheet of paper. In two separate trials participants used their right or their left hands. The authors found that the most substantial deficits in the cognitive tasks (verbal, non-verbal, reading comprehension, and mathematical ability) corresponded to those children that were closer to the point of equal hand skill, exhibiting what they called "hemispheric indecision" (Crow et al., 1998). The authors suggest that failure to establish hemispheric dominance unequivocally is problematic and that lack of dominance by age 11 results in global delays in cognitive development. Supporting this finding, a more recent study showed that children with consistent hand use and superior skill of the preferred hand obtained better scores in reading and mathematics (Cheyne et al., 2010). Other studies, however, have failed to find a relationship between lateralized hand use and cognitive abilities (Mayringer and Wimmer, 2002). Crow et al. (1998) however, suggested that this might be attributed to a failure in appreciating handedness as a continuum rather than an absolute. Our results support this view because rather than considering children as right-handed or left-handed, their hand preference was evaluated by hand use in a natural (unconstrained as to what hand or grip to use) grasping task. In our experiment, all children self-reported as right-handed, yet many of them failed to show a clear right hand preference for grasping. Overall, these children's BRIEF scores indicated more problems with executive function. In other words, our grasping tasks produced a continuum of right hand use rather than an absolute preference that correlated and more importantly, predicted EF. It remains to be shown if left-handed children that display a very strong left hand preference (thus strong right hemisphere lateralization) also enjoy enhanced EF. Regardless of handedness, if the degree of lateralization supports better motor and cognitive performance, then we would predict that very strongly left-handed individuals would show similar advantages to those with a strong right hand preference.

Developmental research has provided evidence that by birth, both anatomical and functional lateralization are features of the human brain (for a recent review see Hervé et al., 2013). Furthermore, studies have shown that compared to other brain circuits, regions subserving motor control are established and refined earlier (Lin et al., 2008; Dubois et al., 2009; Ratnarajah et al., 2013). Ratnarajah et al. used DTI to determine the pattern of structural connectivity asymmetry in 124 normal neonates. Their results showed that the left hemisphere exhibits greater structural efficiency than does the right hemisphere, and they conclude that this early specialized connectivity supports lateralized functional need, particularly in the motor domain. This evidence suggests that anatomical asymmetries exist at birth and functional lateralization continues to mature during childhood (Hervé et al., 2013). Our results are in line with these findings. Children in the older age group displayed greater preference for using their right hand during grasping as well as lower scores on the BRIEF, which indicates better EF. Although speculative, it is possible that greater structural efficiency in the left hemisphere contributes to stronger right hand preference and EF. Clearly this relationship deserves further consideration. Our results suggest the interesting possibility of utilizing measures of motor lateralization for predicting deviations from normal developmental trajectories, specifically for EF. This suggestion would be supported by studies showing the power of using motor skill as a predictor of later cognitive abilities. For example Johnson et al. (1995) showed that fine motor tasks predict kindergarten readiness and other have found correlations between fine motor skills and reading and mathematical achievement (Wolff et al., 1985; Luo et al., 2007). To our knowledge no study has introduced measures of hand and space lateralization as a tool to assess cognitive function, let alone as a means to enhance these processes. We speculate that those studies showing that better fine motor skill correlate with better cognitive abilities might be in part related to the strength of hand preference (i.e., lateralization). It is well-known that proficiency in a manual activity is related to the amount of practice during the learning period (e.g., Jabusch et al., 2009). Furthermore, it has been shown that training-induced brain plasticity after motor sequence learning persists for months (Karni, 1995). We propose that working on hand skills that promote lateralization might be an effective method to enhance EF.

A strength of the current study was the degree to which hand and space use correlated and further predicted the GEC of the BRIEF. Furthermore, both indices and many subscales of the BRIEF correlated with hand and space use. The only subscale that never correlated with any of the grasping measures was *shift*. This is not surprising, as we believe our tasks did not require the child to shift problem-solving strategies to be successful. However, it is important to bear in mind that shift contributes to both the Behavior Regulation Index (BRI) and ultimately the GEC. Both of these measures repeatedly correlated with grasping behavior. The subscales of the BRIEF that appeared most often as significantly correlated with our grasping measures were *inhibit, plan*, and *working memory*. As defined by Gioia and Isquith (2004), "inhibit is the ability to resist or delay an impulse, to appropriately stop one's own activity at the proper time, or both; plan involves anticipating future events, setting goals and developing appropriate steps ahead of time to carry out an associated task or action; working memory is the process of holding information in mind for the purpose of completing a related task." Both grasping tasks demand recruitment of these three components for successful completion. For example, in the grasp-to-construct the child must: (1) resist the impulse of grabbing all the pieces at once, and/or assembling a structure of their own design (*inhibit*), (2) develop the appropriate steps ahead of time to reproduce the sample model (*plan*), and (3) keep in memory the goal of the task (*working memory*). Similarly, in the grasp-to-place task the child must wait, listen to, and follow the instruction as to which blocks to grasp (inhibit, planning, working memory). In both cases a motor plan must be created and executed in order to grasp the blocks. We believe these tasks tested the fundamental essence of the *inhibit, plan*, and *working memory* subscales. Our results align with a trend in the literature which has shown *inhibit*, *plan*, and *working memory* as a reliable measures of EF (Moriguchi and Hiraki, 2013; for reviews see: Jurado and Rosselli, 2007; Best and Miller, 2010).

The results from the regression analyses highlight the interconnectedness of EF and lateralization for hand and space use. To find out which variables were useful as predictors of others, chronological age, right-hand use, right-near space use, and GEC were each used separately as dependent measures. Notably, we found that both hand and space use are predictors of EF. In turn, EF is a predictor of right hand use and space use. In other words, the more children used their right hand or the right near space for grasping, the better their EF scores and vice versa. This is a remarkable finding that could have implications for intervention. There is emerging evidence that motor activity such as aerobic exercise (Hillman et al., 2008; Chaddock et al., 2011), bimanual basketball dribbling (Davis et al., 2011) and handwriting (Rosenblum, 2013) improves aspects of executive function. What remains to be shown is whether short-term motor interventions that promote the use of the right hand during skill grasping have a beneficial effect on EF. The regression analyses also showed a reciprocal relationship between chronological age and right-near space use. The older the child, the more likely they are to grasp in right near space first and vice versa. This result is consistent with our hypothesis that as children age their use of space resembles the adult pattern, that is, right-handed adults prefer to grasp in right-near space followed by equal use of left-near and right-far space (de Bruin et al., 2014). The results suggest that there is a maturation time-line for space use. In light of the current results, these issues warrant further investigation.

A limitation of this study was the exclusive use of the BRIEF as our measure of EF. Clearly additional in-house tests of EF would both inform and complement the assessment of these processes. Future investigations aimed at a more comprehensive assessment of EF might further substantiate the current findings.

In conclusion, the results from the present investigation suggest finer measures that afford an examination of hand and space use preference for grasping should be included to complement existing strategies for early detection of developmental delays, particularly if EF truly predicts school achievement and life success.

ACKNOWLEDGMENTS

This project was funded by grants from the National Science and Engineering Research Council (NSERC Canada) to Claudia L. R. Gonzalez and the Office of Research Services at the University of Lethbridge to Claudia L. R. Gonzalez, Inge Genee, Fangfang Li, Noella Piquette, Nicole Rosen and Robbin Gibb.

REFERENCES

Abe, M., and Hanakawa, T. (2009). Functional coupling underlying motor and cognitive functions of the dorsal premotor cortex. *Behav. Brain Res.* 198, 13–23. doi: 10.1016/j.bbr.2008.10.046

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., and Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain* 135(Pt 4), 1154–1164. doi: 10.1093/brain/aws021

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Butcher, P. R., van Braeckel, K., Bouma, A., Einspieler, C., Stremmelaar, E. F., and Bos, A. F. (2009). The quality of preterm infants' spontaneous movements: an early indicator of intelligence and behaviour at school age. *J. Child Psychol. Psychiatry* 50, 920–930. doi: 10.1111/j.1469-7610.2009.02066.x

Cameron, C. E., Brock, L. L., Murrah, W. M., Bell, L. H., Worzalla, S. L., Grissmer, D., et al. (2012). Fine motor skills and executive function both contribute to kindergarten achievement. *Child Dev.* 83, 1229–1244. doi: 10.1111/j.1467-8624.2012.01768.x

Carlson, A. G., Rowe, E., and Curby, T. W. (2013). Disentangling fine motor skills' relations to academic achievement: the relative contributions of visual-spatial integration and visual-motor coordination. *J. Genet. Psychol.* 174, 514–533. doi: 10.1080/00221325.2012.717122

Chaddock, L., Pontifex, M. B., Hillman, C. H., and Kramer, A. F. (2011). A review of the relation of aerobic fitness and physical activity to brain structure and function in children. *J. Int. Neuropsychol. Soc.* 17, 1–11. doi: 10.1017/S1355617711000567

Cheyne, C. P., Roberts, N., Crow, T. J., Leask, S. J., and Garcia-Finana, M. (2010). The effect of handedness on academic ability: a multivariate linear mixed model approach. *Laterality* 15, 451–464. doi: 10.1080/13576500902976956

Crow, T. J., Crow, L. R., Done, D. J., and Leask, S. (1998). Relative hand skill predicts academic ability: global deficits at the point of hemispheric indecision. *Neuropsychologia* 36, 1275–1282. doi: 10.1016/S0028-3932(98)00039-6

Davis, E. E., Pitchford, N. J., and Limback, E. (2011). The interrelation between cognitive and motor development in typically developing children aged 4-11 years is underpinned by visual processing and fine manual control. *Br. J. Psychol.* 102, 569–584. doi: 10.1111/j.2044-8295.2011.02018.x

de Bruin, N., Bryant, D. C., and Gonzalez, C. L. (2014). "Left neglected," but only in far space: spatial biases in healthy participants revealed in a

visually guided grasping task. *Front. Neurol.* 5:4. doi: 10.3389/fneur.2014.00004

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Dev.* 71, 44–56. doi: 10.1111/1467-8624.00117

Diamond, A. (2007). Interrelated and interdependent. *Dev. Sci.* 10, 152–158. doi: 10.1111/j.1467-7687.2007.00578.x

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529

Donders, J. (2002). The behavior rating inventory of executive function: introduction. *Child Neuropsychol.* 8, 229–230. doi: 10.1076/chin.8.4.229.13508

Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J. F., Le Bihan, D., and Dehaene-Lambertz, G. (2009). Structural asymmetries in the infant language and sensorimotor networks. *Cereb. Cortex* 19, 414–423. doi: 10.1093/cercor/bhn097

Eliason, M. J. (1986). Neurofibromatosis: implications for learning and behavior. *J. Dev. Behav. Pediatr.* 7, 175–179. doi: 10.1097/00004703-198606000-00009

Eliasson, A. C., Rosblad, B., and Forssberg, H. (2004). Disturbances in programming goal-directed arm movements in children with ADHD. *Dev. Med. Child Neurol.* 46, 19–27. doi: 10.1111/j.1469-8749.2004.tb00429.x

Fuentes, C. T., Mostofsky, S. H., and Bastian, A. J. (2009). Children with autism show specific handwriting impairments. *Neurology* 73, 1532–1537. doi: 10.1212/WNL.0b013e3181c0d48c

Gallivan, J. P., McLean, A., and Culham, J. C. (2011). Neuroimaging reveals enhanced activation in a reach-selective brain area for objects located within participants' typical hand workspaces. *Neuropsychologia* 49, 3710–3721. doi: 10.1016/j.neuropsychologia.2011.09.027

Gioia, G. A., and Isquith, P. K. (2004). Ecological assessment of executive function in traumatic brain injury. *Dev. Neuropsychol.* 25, 135–158. doi: 10.1080/87565641.2004.9651925

Gioia, G. A., Isquith, P. K., Guy, S. C., and Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychol.* 6, 235–238. doi: 10.1076/chin.6.3.235.3152

Glascher, J., Rudrauf, D., Colom, R., Paul, L. K., Tranel, D., Damasio, H., et al. (2010). Distributed neural system for general intelligence revealed by lesion mapping. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4705–4709. doi: 10.1073/pnas.0910397107

Gonzalez, C. L., Ganel, T., and Goodale, M. A. (2006). Hemispheric specialization for the visual control of action is independent of handedness. *J. Neurophysiol.* 95, 3496–3501. doi: 10.1152/jn.01187.2005

Gonzalez, C. L., and Goodale, M. A. (2009). Hand preference for precision grasping predicts language lateralization. *Neuropsychologia* 47, 3182–3189. doi: 10.1016/j.neuropsychologia.2009.07.019

Gonzalez, C. L., Whitwell, R. L., Morrissey, B., Ganel, T., and Goodale, M. A. (2007). Left handedness does not extend to visually guided precision grasping. *Exp. Brain Res.* 182, 275–279. doi: 10.1007/s00221-007-1090-1

Hanakawa, T. (2011). Rostral premotor cortex as a gateway between motor and cognitive networks. *Neurosci. Res.* 70, 144–154. doi: 10.1016/j.neures.2011.02.010

Hervé, P. Y., Zago, L., Petit, L., Mazoyer, B., and Tzourio-Mazoyer, N. (2013). Revisiting human hemispheric specialization with neuroimaging. *Trends Cogn. Sci.* 17, 69–80. doi: 10.1016/j.tics.2012.12.004

Hillman, C. H., Erickson, K. I., and Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *Nat. Rev. Neurosci.* 9, 58–65. doi: 10.1038/nrn2298

Iverson, J. M. (2010). Developing language in a developing body: the relationship between motor development and language development. *J. Child Lang.* 37, 229–261. doi: 10.1017/S0305000909990432

Jabusch, H. C., Alpers, H., Kopiez, R., Vauth, H., and Altenmuller, E. (2009). The influence of practice on the development of motor skills in pianists: a longitudinal study in a selected motor task. *Hum. Mov. Sci.* 28, 74–84. doi: 10.1016/j.humov.2008.08.001

Jacquet, A. Y., Esseily, R., Rider, D., and Fagard, J. (2012). Handedness for grasping objects and declarative pointing: a longitudinal study. *Dev. Psychobiol.* 54, 36–46. doi: 10.1002/dev.20572

Johnson, L. J., Gallagher, R. J., Cook, M., and Wong, P. (1995). Critical skills for kindergarten: perceptions From kindergarten teachers. *J. Early Interv.* 19, 315–327. doi: 10.1177/105381519501900406

Jongbloed-Pereboom, M., Nijhuis-van der Sanden, M. W., Saraber-Schiphorst, N., Crajé, C., and Steenbergen, B. (2013). Anticipatory action planning increases from 3 to 10 years of age in typically developing children. *J. Exp. Child Psychol.* 114, 295–305. doi: 10.1016/j.jecp.2012.08.008

Jurado, M. B., and Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychol. Rev.* 17, 213–233. doi: 10.1007/s11065-007-9040-z

Karni, A. (1995). When practice makes perfect. *Lancet* 345, 395. doi: 10.1016/S0140-6736(95)90386-0

Kirby, A., Sugden, D., Beveridge, S., Edwards, L., and Edwards, R. (2008). Dyslexia and developmental co-ordination disorder in further and higher education-similarities and differences. Does the 'label' influence the support given?*Dyslexia* 14, 197–213. doi: 10.1002/dys.367

Kolb, B., and Whishaw, I. Q. (2009). *Fundamentals of Human Neuropsychology, 6th Edn*. New York, NY: Worth Publishers.

Lalonde, G., Henry, M., Drouin-Germain, A., Nolin, P., and Beauchamp, M. H. (2013). Assessment of executive function in adolescence: a comparison of traditional and virtual reality tools. *J. Neurosci. Methods* 219, 76–82. doi: 10.1016/j.jneumeth.2013.07.005

Lin, W., Zhu, Q., Gao, W., Chen, Y., Toh, C. H., Styner, M., et al. (2008). Functional connectivity MR imaging reveals cortical functional connectivity in the developing brain. *Am. J. Neuroradiol.* 29, 1883–1889. doi: 10.3174/ajnr.A1256

Luo, Z., Jose, P. E., Huntsinger, C. S., and Pigott, T. D. (2007). Fine motor skills and mathematics achievement in east asian american and european american kindergartners and first graders. *Br. J. Dev. Psychol.* 25, 595–614. doi: 10.1348/026151007X185329

Masten, A. S., Herbers, J. E., Desjardins, C. D., Cutuli, J. J., McCormick, C. M., and Sapienza, J. K.,. Zelazo, P. D. (2012). Executive function skills and school success in young children experiencing homelessness. *Educ. Res.* 41, 375–384. doi: 10.3102/0013189X12459883

Mayringer, H., and Wimmer, H. (2002). No deficits at the point of hemispheric indecision. *Neuropsychologia* 40, 701–704. doi: 10.1016/S0028-3932(01)00191-9

Michel, G. F., Tyler, A. N., Ferre, C., and Sheu, C. F. (2006). The manifestation of infant hand−use preferences when reaching for objects during the seven−to thirteen−month age period. *Dev. Psychobiol.* 48, 436–443. doi: 10.1002/dev.20161

Momennejad, I., and Haynes, J. D. (2013). Encoding of prospective tasks in the human prefrontal cortex under varying task loads. *J. Neurosci.* 33, 17342–17349. doi: 10.1523/JNEUROSCI.0492-13.2013

Moriguchi, Y., and Hiraki, K. (2013). Prefrontal cortex and executive function in young children: a review of NIRS studies. *Front. Hum. Neurosci.* 7:867. doi: 10.3389/fnhum.2013.00867

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Pangelinan, M. M., Zhang, G., VanMeter, J. W., Clark, J. E., Hatfield, B. D., and Haufler, A. J. (2011). Beyond age and gender: relationships between cortical and subcortical brain volume and cognitive-motor abilities in school-age children. *Neuroimage* 54, 3093–3100. doi: 10.1016/j.neuroimage.2010.11.021

Pennequin, V., Sorel, O., and Fontaine, R. (2010). Motor planning between 4 and 7 years of age: changes linked to executive functions. *Brain Cogn.* 74, 107–111. doi: 10.1016/j.bandc.2010.07.003

Piek, J. P., Dawson, L., Smith, L. M., and Gasson, N. (2008). The role of early fine and gross motor development on later motor and cognitive ability. *Hum. Mov. Sci.* 27, 668–681. doi: 10.1016/j.humov.2007.11.002

Racine, M. B., Majnemer, A., Shevell, M., and Snider, L. (2008). Handwriting performance in children with attention deficit hyperactivity disorder (ADHD). *J. Child Neurol.* 23, 399–406. doi: 10.1177/0883073807309244

Ratnarajah, N., Rifkin-Graboi, A., Fortier, M. V., Chong, Y. S., Kwek, K., Saw, S. M., et al. (2013). Structural connectivity asymmetry in the neonatal brain. *Neuroimage* 75, 187–194. doi: 10.1016/j.neuroimage.2013.02.052

Roebers, C. M., and Kauer, M. (2009). Motor and cognitive control in a normative sample of 7-year-olds. *Dev. Sci.* 12, 175–181. doi: 10.1111/j.1467-7687.2008.00755.x

Rosenbaum, D. A., Chapman, K. M., Weigelt, M., Weiss, D. J., and van der Wel, R. (2012). Cognition, action, and object manipulation. *Psychol. Bull.* 138, 924–946. doi: 10.1037/a0027839

Rosenblum, S. (2013). Handwriting measures as reflectors of executive functions among adults with Developmental Coordination Disorders (DCD). *Front. Psychol.* 4:357. doi: 10.3389/fpsyg.2013.00357

Roth, M., McCaul, E., and Barnes, K. (1993). Who becomes an "at-risk" student? The predictive value of a Kindergarten Screening Battery. *Except Child* 59, 348–358.

Sacrey, L. A., Arnold, B., Whishaw, I. Q., and Gonzalez, C. L. (2013). Precocious hand use preference in reach-to-eat behavior versus manual construction in 1- to 5-year-old children. *Dev. Psychobiol.* 55, 902–911. doi: 10.1002/dev.21083

Sacrey, L. A., Karl, J. M., and Whishaw, I. Q. (2012). Development of rotational movements, hand shaping, and accuracy in advance and withdrawal for the reach-to-eat movement in human infants aged 6-12 months. *Infant Behav. Dev.* 35, 543–560. doi: 10.1016/j.infbeh.2012.05.006

Schmahmann, J. D., and Pandya, D. N. (2008). Disconnection syndromes of basal ganglia, thalamus, and cerebrocerebellar systems. *Cortex* 44, 1037–1066. doi: 10.1016/j.cortex.2008.04.004

Stone, K. D., Bryant, D. C., and Gonzalez, C. L. (2013). Hand use for grasping in a bimanual task: evidence for different roles? *Exp. Brain Res.* 224, 455–467. doi: 10.1007/s00221-012-3325-z

Stone, K. D., and Gonzalez, C. L. (2014). Grasping with the eyes of your hands: hapsis and vision modulate hand preference. *Exp. Brain Res.* 232, 385–393. doi: 10.1007/s00221-013-3746-3

Strauss, E. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary.* New York, NY: Oxford University Press.

Stuss, D. T. (2011). Traumatic brain injury: relation to executive dysfunction and the frontal lobes. [Research Support, Non-U.S. Gov't Review]. *Curr. Opin. Neurol.* 24, 584–589. doi: 10.1097/WCO.0b013e32834c7eb9

Sugden, D., Kirby, A., and Dunford, C. (2008). Issues surrounding children with developmental coordination disorder. *Int. J. Disabil. Dev. Educ.* 55, 173–187. doi: 10.1080/10349120802033691

Thibaut, J. P., and Toussaint, L. (2010). Developing motor planning over ages. *J. Exp. Child Psychol.* 105, 116–129. doi: 10.1016/j.jecp.2009.10.003

van der Meulen, M., Allali, G., Rieger, S. W., Assal, F., and Vuilleumier, P. (2012). The influence of individual motor imagery ability on cerebral recruitment during gait imagery. *Hum. Brain Mapp.* 27, 22192. doi: 10.1002/hbm.22192

van Swieten, L. M., van Bergen, E., Williams, J. H., Wilson, A. D., Plumb, M. S., Kent, S. W., et al. (2010). A test of motor (not executive) planning in developmental coordination disorder and autism. *J. Exp. Psychol.* 36, 493–499. doi: 10.1037/a0017177

Waber, D. P., Gerber, E. B., Turcios, V. Y., Wagner, E. R., and Forbes, P. W. (2006). Executive functions and performance on high-stakes testing in children from urban schools. *Dev. Neuropsychol.* 29, 459–477. doi: 10.1207/s15326942dn2903_5

Willcutt, E. G., and Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder differences by gender and subtype. *J. Learn. Disabil.* 33, 179–191. doi: 10.1177/002221940003300206

Wolff, P. H., Gunnoe, C., and Cohen, C. (1985). Neuromotor maturation and psychological performance: a developmental study. *Dev. Med. Child Neurol.* 27, 344–354. doi: 10.1111/j.1469-8749.1985.tb04546.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 December 2013; accepted: 18 March 2014; published online: 07 April 2014.

Citation: Gonzalez CLR, Mills KJ, Genee I, Li F, Piquette N, Rosen N and Gibb R (2014) Getting the right grasp on executive function. Front. Psychol. ***5****:285. doi: 10.3389/fpsyg.2014.00285*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Gonzalez, Mills, Genee, Li, Piquette, Rosen and Gibb. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 May 2014
doi: 10.3389/fpsyg.2014.00405

The role of executive function in children's source monitoring with varying retrieval strategies

Becky Earhart and Kim P. Roberts*

Department of Psychology, Wilfrid Laurier University, Waterloo, ON, Canada

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Yasuhiro Kanakogi, Kyoto University, Japan
Mika Naito, Joetsu University of Education, Japan

***Correspondence:**
Kim P. Roberts, Department of Psychology, Wilfrid Laurier University, 75 University Avenue West, Waterloo, ON N2L 3C5, Canada
e-mail: kroberts@wlu.ca

Previous research on the relationship between executive function and source monitoring in young children has been inconclusive, with studies finding conflicting results about whether working memory and inhibitory control are related to source-monitoring ability. In this study, the role of working memory and inhibitory control in recognition memory and source monitoring with two different retrieval strategies were examined. Children ($N = 263$) aged 4–8 participated in science activities with two sources. They were later given a recognition and source-monitoring test, and completed measures of working memory and inhibitory control. During the source-monitoring test, half of the participants were asked about sources serially (one after the other) whereas the other half of the children were asked about sources in parallel (considering both sources simultaneously). Results demonstrated that working memory was a predictor of source-monitoring accuracy in both conditions, but inhibitory control was only related to source accuracy in the parallel condition. When age was controlled these relationships were no longer significant, suggesting that a more general cognitive development factor is a stronger predictor of source monitoring than executive function alone. Interestingly, the children aged 4–6 years made more accurate source decisions in the parallel condition than in the serial condition. The older children (aged 7–8) were overall more accurate than the younger children, and their accuracy did not differ as a function of interview condition. Suggestions are provided to guide further research in this area that will clarify the diverse results of previous studies examining whether executive function is a cognitive prerequisite for effective source monitoring.

Keywords: executive function, cognitive development, source monitoring, working memory, inhibitory control

INTRODUCTION

Between the ages of 3- and 8-years-old many fundamental changes in cognitive development occur rapidly, and one cognitive skill that shows drastic improvement is source monitoring (Roberts, 2002). The term "source" refers to the conditions under which a memory was acquired including many different attributes, such as when or where an event occurred or how it was perceived (e.g., whether it actually happened or was only imagined or seen on television; Johnson et al., 1993). Source monitoring is the process of making decisions about the source of a memory.

Although there is much research establishing that young children do not perform as well on source-monitoring tasks as older children or adults do, it is less clear why these age differences occur or what cognitive changes are supporting the development of source monitoring during early childhood. Thus, researchers are now addressing not only what recall strategies are most successful with young children, but also what underlying cognitive processes contribute to the development of source monitoring (e.g., Ruffman et al., 2001; Roberts and Powell, 2005; Kanakogi et al., 2012). Executive function is vital for many cognitive abilities, and this study addresses whether different aspects of executive function are necessary for the development of source monitoring in early childhood, or whether executive function is only more generally related to episodic memory. In particular, the ability to recall one source in the face of competing sources and to "compare and contrast" sources may be related to accurate source monitoring. The proposed executive function-source monitoring relationship was tested using two variations of a source-monitoring task: children were either asked to recall one source at a time *or* to consider two sources simultaneously in order to provide information about how executive function processes might contribute in somewhat different ways in each of these conditions.

THE SOURCE-MONITORING FRAMEWORK

The Source-Monitoring Framework (Johnson et al., 1993) was developed to explain how source judgments are made. According to the framework, making attributions about the origin of a memory is a complex decision-making process that is more complicated than simple retrieval of source information because one can remember an event but not the circumstances under which the event occurred (e.g., who spoke, whether it was a dream). Source decisions are often based on the qualitative characteristics of memory traces, such as the spatial or temporal context, the amount of perceptual detail, the cognitive operations associated with the memory, semantic details, and the affective response when the memory was formed. Critically, memories originating

from the self (e.g., dreams, own actions, thinking through a plan) comprise a different qualitative profile than memories of external sources such as a colleague describing an issue to you or watching a movie. Typically memories of internal (self-generated) events contain fewer perceptual details and more cognitive operations than memories of events derived external to an individual; the profile of externally-derived memories is the reverse of internal sources. For example, when distinguishing between an event that actually happened and something that was only imagined, it could be expected that a real event would have greater perceptual detail and fewer records of cognitive operations than an imagined event.

Source decisions are made either through heuristic or systematic judgment processes. Heuristic processes involve quick decisions that may occur in the course of remembering without conscious awareness of making a decision (e.g., remembering the source of a memory because you recalled it in the person's voice; Johnson et al., 1993). Systematic judgment processes are more analytic and deliberate; when making systematic decisions, a person will reason carefully about what is possible given the information that they have from the memory itself. This may involve retrieving supporting memories, reasoning about constraints, and employing strategies (Johnson et al., 1993).

THE DEVELOPMENT OF SOURCE MONITORING

As mentioned earlier, extensive past research on source monitoring has clearly established that young children (aged 3–4 years) have substantially poorer source-monitoring abilities than do older children (e.g., 8–12 year-olds) or adults (e.g., Gopnik and Graf, 1988; Lindsay et al., 1991; Powell and Thomson, 1996; Roberts and Blades, 1999). It is not until approximately age 10 that children perform as well as adults on many source-monitoring tasks (see Roberts, 2002, for a review). Given that the quality of the memory and the quality of the decision-making process are both important components of accurate source monitoring (Johnson et al., 1993), young children may not have the necessary cognitive prerequisites to engage in systematic processing. For example, young children may not be able to coordinate the many decision-making processes necessary for systematic processing, and therefore cannot reason about the constraints of memories in order to problem-solve. The maturational development of the frontal lobe may have to be in motion before complex and effortful decisions can be made (De Luca and Leventer, 2008).

EXECUTIVE FUNCTION AND SOURCE-MONITORING DECISIONS

Research on adults with frontal lobe damage demonstrates that these participants show deficits in memory and fail to monitor the sources of their memories effectively. Adults with frontal lobe damage show many of the same problems that young children demonstrate in source monitoring, suggesting that the frontal lobe is implicated in the development of source monitoring (Schacter et al., 1995). Differences in executive function, a broad category of skills that support goal-directed behavior, have been linked with immature frontal lobe development (e.g., De Luca and Leventer, 2008). Executive function underlies many cognitive abilities, and can potentially be linked with skills required for source monitoring.

There are theoretical reasons to believe that two components of executive function, inhibitory control and working memory, would be related to source-monitoring accuracy. Inhibitory control might be related because source monitoring requires inhibition of the familiarity-based retrieval processes that are often used automatically to make recognition decisions (Ruffman et al., 2001), as well as inhibition of information from competing sources. Working memory may be related to source monitoring because it is involved in controlling attention, and therefore plays a role in designating what information cognitive resources will be allotted to (Gerrie and Garry, 2007). A complex process of reasoning about the constraints of memories, retrieving supporting memories, comparing and contrasting sources, and inhibiting competing information may be needed to make effective decisions about source.

Research in this area has looked for links between executive function and source monitoring, as well as susceptibility to suggestibility and false memories (which are source-monitoring errors in which people fail to recognize the source of information that was suggested to them). A study examining this relationship in an elderly population found that those who had high composite scores on a battery of executive function tests were better at a source-monitoring task involving voice identification than a group who scored low on the executive function tests (Glisky et al., 1995). This suggests that impairments in executive function may partially account for source-monitoring decline in old age.

Research with younger adults has shown connections between working memory and resistance to misleading suggestions. Using a misinformation paradigm where participants were exposed to misleading post-event information, Jaschinski and Wentura (2002) found that adults with a higher working memory capacity were misled to a lesser extent by misinformation than adults with a lower working memory capacity. Gerrie and Garry (2007) replicated these findings and added that the effect of working memory capacity was especially strong for crucial event details as opposed to non-crucial, or peripheral, details. Researchers have suggested that working memory is negatively related to susceptibility to false memories because people with higher working memory capacities are better able to monitor the sources of information (Gerrie and Garry, 2007), and this was supported by a study that showed source monitoring as a mediator between working memory and false recall (Unsworth and Brewer, 2010).

Studies addressing the relationship between executive function and source monitoring in children are more informative about whether executive function underlies developmental improvements in source-monitoring accuracy. However, studies specifically examining this relationship with children are not extensive, and have contradictory results. Several studies have shown that executive function is not predictive of source-monitoring accuracy in children. For example, one study found that cognitive shifting predicted source monitoring, but inhibitory control did not (Kanakogi et al., 2012). In a comprehensive review of individual difference factors in suggestibility, Bruck and Melnyk (2004) reported that there was typically a negative relationship between executive function and suggestibility, but that few studies showed significant correlations.

Other studies, on the other hand, have supported the role of executive function in source-monitoring development. Roberts and Powell (2005) found that children with better inhibitory control were less suggestible to misleading information, and Karpinski and Scullin (2009) found that preschoolers with better executive function were less suggestible in a pressured suggestive interview. The latter study showed significant relationships for both inhibitory control and working memory.

Melinder et al. (2006) found mixed results, with inhibitory control as a significant predictor of suggestibility, but not source monitoring. Similarly, Ruffman et al. (2001) have found a relationship between inhibitory control and source accuracy, but only for some types of source-monitoring questions. They did, however, find clear evidence that working memory was related to source monitoring. Overall, the results of this literature are inconclusive with some researchers finding significant relationships and others finding no evidence of the role of executive function.

In addition, several studies finding significant relationships between suggestibility or source monitoring and executive function have used memory tasks that load heavily on both recognition and source memory. For example, in a typical suggestibility paradigm, children are asked whether or not suggested details occurred during a real event; this requires a simultaneous assessment of whether the child recognizes the detail, as well as its source (real or suggested; e.g., Roberts and Powell, 2005; Karpinski and Scullin, 2009). Similarly, in source-monitoring studies children may be asked whether details occurred in source A, source B, or neither source. Because children are also asked about things that happened in neither source, they must assess both if the detail is familiar, and if so, which source it is from (e.g., Foley et al., 1983; Foley and Johnson, 1985; Ruffman et al., 2001). In these studies, then, it is unclear whether executive function is significantly related to recognition memory, source monitoring, or both.

The present study seeks to add to the body of literature on the relationship between executive function and children's developing source-monitoring skills to test whether executive function is an important predictor of source-monitoring accuracy, or whether executive function plays a more general role in episodic memory after exposure to multiple sources. Our procedure separated recognition and source-monitoring tasks with a two-step test; first children identified the details they had seen, and then they made source decisions about them. With this procedure we sought to look at the relationships of executive function to recognition memory and source monitoring separately, to determine whether executive function makes a unique contribution to source monitoring when recognition memory demands are removed.

SERIAL vs. PARALLEL RECALL OF SOURCES

The cognitive abilities involved in source monitoring may differ depending on retrieval strategy. In previous studies of source monitoring children have been asked about sources serially (one source at a time, reporting everything they remember about one source, followed by everything they remember about another source; Thierry et al., 2001, 2005) or in parallel (questioned about multiple sources of information simultaneously; Powell and Thomson, 2003). The role of executive function in source monitoring may vary depending on the way the task is structured.

If children are asked to consider sources serially, they are required to inhibit the reporting of information from other irrelevant sources, including sources within the same event (Roberts and Powell, 2005). Consequently, inhibitory control may play a stronger role in source monitoring when the task involves reporting information about one source at a time. On the other hand, when children are asked to consider sources in parallel, they must hold information about the characteristics of multiple sources in working memory in order to compare and contrast them. When asked about multiple sources at the same time working memory may play a more important role in source-monitoring ability. Therefore, this study also tested whether inhibitory control and working memory would be differentially related to source-monitoring accuracy when children are questioned about sources serially vs. in parallel, because recalling information from one source while holding back information from other sources may require different cognitive skills than comparing and contrasting competing sources.

Additionally, different retrieval strategies may lead to differences in source-monitoring accuracy. Source judgments are often based on comparing the relative strength of characteristics (e.g., the amount of perceptual detail) to determine which source "fits" better with a memory (Johnson et al., 1993). The Source-Monitoring Framework might predict that source decisions would be more accurate when considering multiple sources at the same time, because thinking about different sources at the same time would enable a more direct comparison of relevant characteristics than thinking about sources in a serial fashion. This could lead to higher source-monitoring accuracy compared to asking about sources serially, when a strategy is not facilitated. If asked about sources serially, children would be required to spontaneously generate the comparison strategy in order to source monitor with similar accuracy levels. A third goal of this study was to empirically test the prediction that source-monitoring accuracy would be higher for children who were asked about sources in parallel than those asked about sources serially.

THE PRESENT STUDY

This study examined the relationship between executive function, recognition memory and source monitoring both generally and with two different retrieval strategies (serial vs. parallel). Children aged 4–8 participated in science activities interactively and by listening to a story (i.e., the target sources). The children were given a recognition and source-monitoring test after four to seven days. During a third session, children's working memory and inhibitory control were measured to determine whether these cognitive variables were related to source monitoring, and whether the role of these two variables differed for children in the serial and parallel conditions. Following from the discussion above, this study had four hypotheses:

Hypothesis 1: Age differences were expected in executive function, recognition memory and source monitoring, consistent with the large body of literature demonstrating

development across childhood in all of these areas. Though this hypothesis was not novel, we wanted to confirm that there were in fact age differences in both executive function and memory before attempting to examine relationships between them.

Hypothesis 2: It was expected that working memory and inhibitory control would be significant predictors of both recognition and source memory.

Hypothesis 3: An interaction was predicted such that working memory would be a stronger predictor of source accuracy in the parallel condition, whereas inhibitory control would be a stronger predictor of source accuracy in the serial condition.

Hypothesis 4: Overall differences in source-monitoring accuracy between the serial and parallel interview conditions were predicted, with children making more accurate source decisions when considering sources in parallel than when considering sources serially.

METHODS AND MATERIALS

DESIGN

This study had a 3 (Age in years: 4–5, 6, 7–8) × 2 (Interview Condition: Serial, Parallel) × 2 (Source Presentation: Real-Life, Story) mixed design with the last factor within-subjects.

PARTICIPANTS

Initially 308 children from local daycares, elementary schools, and a university summer day camp participated. Eighteen participants did not complete the study because they missed a session and an additional 27 who completed the study were excluded [12 did not provide any details in free recall, indicating that they did not remember the activities; 3 showed evidence of a yes bias (i.e., a response bias toward saying yes they recognized every detail, including misleading ones) and 12 were excluded due to interviewer errors (e.g., asking source questions about details the child said were not present at the activities)]. The excluded participants were equally distributed across age groups and interview conditions.

The final sample was 263 4- to 8-year-old children (52% male). Four participants did not complete the executive function tests due to time constraints during testing and therefore were excluded from analyses of the cognitive variables, but their memory scores were still included in source accuracy comparisons between age groups and interview conditions. The 4- to 5-year-olds ($n = 84$, 41 in the serial condition) had a mean age of 5.04 ($SD = 0.59$), the 6-year-olds ($n = 79$, 40 in the serial condition) had a mean age of 6.46 ($SD = 0.28$) and the 7- to 8-year-olds ($n = 100$, 49 in the serial condition) had a mean age of 7.87 ($SD = 0.59$). The children were recruited from a mid-sized Canadian city. Information about participants' ethnicity was not available, but the majority of participants were Caucasian and from middle-class families. Informed consent was obtained from a parent/guardian prior to the beginning of data collection, and children assented to participate. There was no monetary compensation for participation. Participants were randomly assigned to one of two interview conditions with the constraint that there were approximately equal numbers of children from each age group and gender in each interview condition.

MATERIALS AND PROCEDURE

Event

Groups of up to 10 children participated in science activities about the human body comprising an interactive activity referred to as the "real-life demonstration" and a story. There were two presentation scripts and they were counterbalanced so that each was shown as the real-life demonstration half of the time and the story the other half of the time. The order of the presentations was also counterbalanced. The presentations each lasted approximately 10 min and had similar content (i.e., a researcher using simple experiments and science materials to teach children about the body) but the research assistant that conducted the demonstration and the main character of the story were different people. The story was a PowerPoint presentation with text and photos presented on a laptop.

The sources were clearly labeled for the children throughout the event by repeatedly referring to them as "the real-life demonstration" and "the story." In each presentation there were 12 details that would be tested during the memory interview, and these details were highlighted during the presentations to ensure that children paid attention to and encoded them.

Baseline memory test

A baseline memory assessment was administered immediately after the event to measure encoding. The relationship between this measure of recognition memory after no delay and executive function was tested, but this measure also served to ensure that there were no differences between interview conditions in initial event memory. The test included 10 recognition questions about event details that were not included in the later memory interview (five from the real-life demonstration and five from the story). The questions were asked in random order, with no reference to the source of the details. Accuracy proportions were calculated for analyses by dividing the number of correct answers by the number of questions asked.

Memory interview

After four to seven days children were interviewed individually by a new research assistant who was blind to counterbalancing condition and therefore was not aware of whether the children were correct or not when choosing the source of details. At the beginning of the 30-min interview, the interviewer introduced herself and spent a few minutes building rapport with the child. The children were given the chance to freely recall anything they could remember about the activities in response to open-ended prompts about what happened, confirming that they remembered the activities and both sources. When children had reported everything they could remember, the interviewer continued to the recognition and source questions.

The memory test was a modified version of the posting-box procedure (Bright-Paul et al., 2005). Participants were required to sort cards depicting details into boxes that represented the sources of the details. Ideally, children would have high hit rates so that they could make source judgments about many event details that they experienced. As well, for source-monitoring accuracy, the ideal level of task difficulty would be such that children were performing above chance but not at ceiling in order that the

sample would have enough variability to be related to other variables. This procedure yielded optimal results with children doing well on source monitoring, but not performing at ceiling in any age group (see the Results section Source Accuracy for further details).

There were 36 photographs (3 × 4 inches) comprising the 12 non-misleading details from each source and 12 misleading details that were not presented in either source. Children in both conditions first completed a *recognition task*. They were asked to place pictures of details that they remembered from the event in a "Yes" box, and pictures they did not remember from the event in a "No" box (a "Don't Know" box was also available). The cards were shuffled and shown to children one at a time as the interviewer asked about the details (e.g., "Did you ever see *dirt from the garden* at the activities?"). Once children had sorted all 36 cards, the interviewer took cards placed in the "Don't Know" box and gave children a second opportunity to sort through those cards before asking source questions.

The subsequent *source-monitoring task* began by retrieving the cards from the "Yes" box (i.e., details children claimed were in the event). Children were asked to sort the cards into three different boxes to indicate their source: "Real-Life Demonstration," "Story," and "Don't Know." However, these boxes were presented differently to children in the serial and parallel conditions.

For the children in the *serial condition*, the interviewer presented one box at a time (order was counterbalanced) and thus, children were required to consider the sources one after the other. Cards were laid out four at a time and the interviewer provided a label for each picture. The children were asked to look through the cards carefully and put any pictures from the story (for those children with the story box first) in the "Story" box. The remaining cards were set aside. After going through all the cards, the interviewer presented the children with the other source box, and the children went through the leftover cards again, now considering the second source (e.g., the "Real-Life Demonstration" box). Any cards that were not attributed to either source were recorded as "Don't Know."

For the children in the *parallel condition*, the interviewer brought out the "Story" box, the "Real-Life Demonstration" box and the "Don't Know" box at the same time. The interviewer showed the children cards one at a time and labeled the picture, and as the children considered each detail they decided if it belonged in the "Story" box or the "Real-Life Demonstration" box (or the "Don't Know" box if they were unsure about the source). In this condition, the children considered both sources as they thought about where they saw each detail because they had to decide whether it came from the story or the real-life demonstration. After completing the interview, children were thanked for their participation and brought back to their classrooms.

Proportions were calculated for hits (correct identification of non-misleading details), false alarms (incorrect identification of misleading details as having been present at the activities) and source accuracy for story details and real-life details separately. A recognition accuracy score was then calculated by subtracting the proportion of false alarms from the proportion of hits. "Don't know" responses were conservatively coded as incorrect for both recognition and source scores. Scores were summed by two independent coders to prevent errors. The nature of the coding was very objective (i.e., counting correct responses), so inter-rater reliability was greater than 99%. The few disagreements were due to addition errors and were resolved before data analysis.

Cognitive assessments

Within approximately one week of the interview, participants completed a third session individually for approximately 15 min. Children were given a battery of cognitive tests consisting of two working memory tasks and two inhibitory control tasks. These tasks were presented as games to the children.

Working memory. The working memory tests were from the WISC-IV Digit Span subtest (Wechsler, 2003). In the Forward Digit Span test, the participant heard a sequence of numbers and was asked to repeat the sequence. The first trial began with a sequence of two digits and the sequences got progressively longer, up to a maximum of nine digits. There were two trials for each sequence length, and after successful repetition of at least one of those sequences, the sequence length increased by one digit. Testing continued until the participant failed both trials of a sequence length. Children were scored one point for each correct repetition for a maximum score of 16.

The Backward Digit Span test was conducted similarly to the Forward Digit Span test, but in this task participants heard a sequence of numbers and had to repeat the sequence in backwards order. Again, the test began with sequences of two digits and the sequences increased in length by one digit every two trials. The maximum number of digits in a sequence was eight. Participants were given one example and one practice trial before testing commenced. If they answered the practice trial correctly, testing began and continued until the participants incorrectly answered both trials of a sequence length. If the participants did not answer the practice trial correctly, they were given up to two more practice trials. If they still could not answer correctly, testing was discontinued. Children were scored one point for each correct repetition for a maximum score of 16.

Inhibitory control. Participants completed two measures of inhibitory control that have frequently been used in previous literature and are easy to administer: Luria's Hand Game (e.g., Hughes, 1996; Fahie and Symons, 2003) and the Day/Night Stroop task (e.g., Gerstadt et al., 1994; Reck and Hund, 2011). In Luria's Hand Game, the researcher either pointed a finger or made a fist, and the child was asked to make the opposite hand gesture from what the researcher did (e.g., make a fist when she pointed a finger). There were 20 trials in one randomized order: "Fist, Finger, Finger, Fist, Fist, Finger, Finger, Fist, Finger, Finger, Fist, Fist, Fist, Finger, Fist, Finger, Finger, Fist, Fist, Finger." Children were encouraged to respond as quickly as they could. Participants were given a practice trial of each gesture before beginning. On each trial participants were scored one point if they produced the opposite hand gesture or immediately self-corrected their action. A score out of 20 was computed based on the number of successful trials.

The Day/Night Stroop task is a modified Stroop task for children that involves looking at pictures of day and night and saying the opposite of what the picture represents. The pictures were shown in a PowerPoint presentation on a laptop. These pictures are universally recognizable, even for young children: the "day" picture was a blue sky with a sun and clouds, and the "night" picture was a black sky with a moon and stars. Participants were encouraged to respond as quickly as possible, and the slide was advanced to a new picture as soon as they responded. There were 20 trials in one randomized order: "Night, Day, Night, Night, Day, Night, Day, Day, Night, Day, Day, Night, Night, Day, Night, Day, Day, Night, Day, Night." Children were scored one point for each trial where they said the opposite of what was shown (e.g., saying "day" when shown the picture of a moon), and a score out of 20 was calculated.

The order of the four tests was randomized with the constraint that participants received the Forward Digit Span task prior to the Backward Digit Span task.

RESULTS

ANALYTIC STRATEGY

Analyses were conducted to first explore developmental differences in memory accuracy and executive function, as well as differences in source accuracy between interview conditions. We then analyzed the relationship between memory accuracy and executive function, and whether there was an interaction with interview condition. An alpha level of 0.05 was used to determine significance for all analyses, unless otherwise noted.

PRELIMINARY ANALYSES

Preliminary analyses confirmed that there were no overall differences between interview conditions in age, delay measured in days, baseline memory scores, working memory scores, inhibitory control scores, the number of details freely recalled at the beginning of the interview, or recognition accuracy during the memory test (all $ts \leq 1.57$, all $ps \geq 0.12$). There were also no differences between age groups in delay, $F_{(2,\ 260)} = 2.17$, $p = 0.12$.

All four cognitive scores (two inhibitory control scores and two working memory scores) were significantly correlated with each other, $rs \geq 0.20$, $ps < 0.001$, and when age was controlled the results were similar (although the magnitude of the correlations were smaller); see **Table 1** for the correlations. Although the correlations between the two working memory scores and between the two inhibitory control scores were significant, the magnitude of the correlations did not justify combining the measures into two composite scores. Therefore, analyses were conducted on all four cognitive variables.

Table 1 | Correlations between scores on the inhibitory control tasks and working memory tasks.

n = 259	1	2	3	4
1. Day/night stroop	–	0.25**	0.23**	0.42**
2. Luria's hand game	0.22**	–	0.24**	0.20**
3. WISC digit forward score	0.08	0.21**	–	0.51**
4. WISC digit backward score	0.28**	0.15*	0.35**	–

Partial correlations controlling for age are shown in the bottom half of the table.
**Significant at the 0.05 level (2-tailed).*
***Significant at the 0.01 level (2-tailed).*

Table 2 | Mean number of accurate responses for executive function measures by age group.

Age group (years)	Inhibitory control tasks		Working memory tasks	
	Luria's hand game (maximum 20)	Day/night stroop (maximum 20)	WISC digit span forward (maximum 16)	WISC digit span backward (maximum 16)
4–5 (*n* = 81)	15.60 (2.71)	16.62 (3.50)	6.12 (1.71)	3.53 (1.91)
6 (*n* = 79)	16.39 (2.33)	18.33 (2.09)	7.13 (1.70)	5.00 (1.50)
7–8 (*n* = 99)	16.22 (2.16)	18.63 (1.66)	7.89 (1.48)	5.80 (1.29)
Total (*n* = 259)	16.08 (2.41)	17.90 (2.64)	7.10 (1.78)	4.85 (1.83)

Standard deviations are in parentheses.

DEVELOPMENTAL AND INTERVIEW CONDITION DIFFERENCES

Executive function

The inhibitory control measures showed some evidence of a ceiling effect. There was enough variability, however, to find significant correlations with other variables (see below). The working memory scores showed more variability. The means and standard deviations for the four measures by age group are displayed in **Table 2**. Because all of the cognitive variables were correlated, a one-way multiple analysis of variance (MANOVA) was used to compare the scores from three age groups on all four cognitive variables. There was a significant multivariate effect [Wilk's $\lambda = 0.68$, $F_{(8,\ 506)} = 13.60$, $p < 0.001$, $\eta_p^2 = 0.18$]. Follow-up 3 (Age) One-Way analyses of variance (ANOVAs) showed age differences for all variables ($Fs \geq 2.43$, $ps \leq 0.04$, one-tailed). *Post-hoc* Bonferroni comparisons for the inhibitory control tasks showed that 4- to 5-year-olds had lower inhibitory control scores than 6-year-olds or 7- to 8-year-olds, but the older age groups did not differ. Bonferroni comparisons for the working memory variables showed that all three age groups were different from each other on both measures, demonstrating significant improvements in working memory for each age group.

Recognition accuracy

Across age groups, the proportion of accurate responses (hits and correct rejections) had a mean of 0.81 ($SD = 0.10$), and ranged from 0.47 to 1.00. Recognition scores calculated by subtracting false alarms from hits were subjected to a 3 (Age in years: 4–5, 6, 7–8) One-Way ANOVA to determine whether there were age differences. There was a main effect of age, $F_{(2,\ 260)} = 24.06$, $p < 0.001$, $\eta_p^2 = 0.16$, and Bonferroni *post-hoc* comparisons revealed that all three age groups differed from each other ($M_{4-5} = 0.57$, $SD = 0.21$; $M_6 = 0.64$, $SD = 0.19$; $M_{7-8} = 0.75$, $SD = 0.14$), demonstrating a steady improvement in recognition memory with age.

Source accuracy

The mean source accuracy proportion was 0.71 ($SD = 0.18$), and scores ranged from 0 to 1.00. A 3 (Age: 4–5, 6, 7–8) × 2 (Interview Condition: Serial, Parallel) × 2 (Source Presentation: Real-Life, Story) analysis of covariance (ANCOVA) was run with repeated measures on the last factor to evaluate hypotheses 1 and 4: whether there were developmental differences and/or interview condition differences in source accuracy. The baseline accuracy proportion was included as a covariate because the baseline scores were correlated with age, and it was significant, $F_{(1,\ 256)} = 17.85$, $p < 0.001$, $\eta_p^2 = 0.065$.

The analysis revealed a main effect of age $F_{(2,\ 256)} = 5.18$, $p = 0.006$, $\eta_p^2 = 0.039$, confirming developmental differences in source accuracy. Bonferroni *post-hoc* comparisons showed that the 4- to 5-year-olds ($M = 0.66$, $SD = 0.22$) made fewer accurate source judgments than the 6-year-olds ($M = 0.74$, $SD = 0.17$) or 7- to 8-year-olds ($M = 0.79$, $SD = 0.14$), who did not differ from each other. Even the youngest age group performed above chance (0.50), $t_{(83)} = 6.43$, $p < 0.001$.

There was also a main effect of interview condition, $F_{(1,\ 256)} = 25.72$, $p < 0.001$, $\eta_p^2 = 0.091$ and an age by condition interaction, $F_{(2,\ 256)} = 4.50$, $p = 0.01$, $\eta_p^2 = 0.034$. Children in the parallel condition ($M = 0.76$, $SD = 0.16$) were more accurate than those in the serial condition ($M = 0.66$, $SD = 0.20$). Follow-up t-tests comparing the accuracy scores of children in the serial and parallel conditions within each age group revealed that the condition effect was significant for the 4- to 5-year-old and 6-year-old age groups, $ts \geq -2.77$, $ps \leq 0.007$, but not for the 7- to 8-year-olds, $t_{(98)} = -0.99$, $p = 0.32$.

Finally, there was a main effect of source presentation, $F_{(1,\ 256)} = 6.10$, $p = 0.01$, $\eta_p^2 = 0.023$, but no interactions involving source presentation, $Fs \leq 0.44$, $ps \geq 0.60$. Children made more accurate source judgments about details from the real-life demonstration ($M = 0.83$, $SD = 0.23$) than about details from the story ($M = 0.59$, $SD = 0.30$). See **Table 3** for the mean source accuracy scores by age group, source presentation and interview condition.

Table 3 | Mean source accuracy proportions by age, interview condition and source presentation.

Age group	Interview condition	Source presentation	
		Real-life	Story
4–5	Serial	0.66	0.45
	($n = 41$)	(0.35)	(0.35)
	Parallel	0.80	0.60
	($n = 43$)	(0.21)	(0.28)
6	Serial	0.78	0.54
	($n = 40$)	(0.25)	(0.29)
	Parallel	0.92	0.66
	($n = 39$)	(0.16)	(0.27)
7–8	Serial	0.89	0.64
	($n = 49$)	(0.15)	(0.29)
	Parallel	0.93	0.65
	($n = 51$)	(0.08)	(0.28)

Standard deviations are in parentheses.

THE RELATIONSHIP BETWEEN EXECUTIVE FUNCTION AND MEMORY VARIABLES

The relationship between executive function and memory in the overall sample

As a first step to explore the contribution of executive function in memory and source monitoring, correlations were run among the cognitive variables and all memory tasks in the study. All four of the cognitive variables were significantly correlated with baseline, recognition, and source accuracy scores (rs ranging from 0.13 to 0.41), except that Luria's Hand Game was not correlated with source accuracy, $r_{(257)} = 0.08$, $p = 0.18$ (see **Table 4** for the full set of correlations). Overall there was evidence that the cognitive variables were related to both recognition and source accuracy. However, when partial correlations were run controlling for age, only Stroop and WISC Backward scores were related to baseline and recognition accuracy, and none of the tasks were related to source accuracy.

Linear regression analyses were run to determine whether executive function scores were predictive of baseline, recognition, and source accuracy scores. All four cognitive scores were entered as predictors simultaneously. For baseline accuracy proportion, Luria's Hand Game did not significantly contribute to the variance in the model, but Stroop scores and both WISC Digit Span scores were significant predictors. Therefore, both inhibitory control and working memory were predictive of memory for the event details immediately afterwards. The model accounted for 22% of the variance in baseline memory. When age was added as a predictor only the Stroop scores remained significant. Standardized regression coefficients and their associated test statistics for significant predictors can be found in **Table 5** for all regression analyses reported.

Only one score emerged as a significant and independent predictor of recognition accuracy: the WISC Backward Digit Span. The other three tests did not reach significance, although the significance level for the Stroop task was marginal. The model accounted for 14% of the total variance in recognition accuracy. There was evidence for the role of working memory, but when age was entered into the regression, the WISC Backward scores were only marginally significant as a predictor.

Table 4 | Correlations between memory scores and executive function scores.

***n* = 259**	**Inhibitory control tasks**		**Working memory tasks**	
Memory scores	**Luria's hand game**	**Day/night stroop**	**WISC digit forward**	**WISC digit backward**
Baseline accuracy	0.14*	0.35**	0.31**	0.41**
	(0.09)	(0.22**)	(0.11)	(0.18**)
Recognition accuracy	0.13*	0.25**	0.23**	0.35**
	(0.09)	(0.12)	(0.05)	(0.15*)
Source accuracy	0.08	0.16**	0.23**	0.27**
	(0.04)	(0.05)	(0.10)	(0.10)

Partial correlations controlling for age are in parentheses.

**Significant at the 0.05 level (2-tailed).*

***Significant at the 0.01 level (2-tailed).*

Table 5 | Standardized regression coefficients and test statistics for significant and marginal predictors in regression analyses.

Dependent variable	Predictor	Beta	*t*	*p*	*p* when age is entered
Baseline accuracy	Day/night stroop	0.21	3.45	0.001	0.007
	WISC digit span forward	0.13	2.02	0.04	0.40
	WISC digit span backward	0.25	3.68	<0.001	0.11
Recognition accuracy	Day/night stroop$^{\Delta}$	0.12	1.81	0.07	0.26
	WISC digit span backward	0.26	3.59	<0.001	0.08
Source accuracy	WISC digit span forward$^{\Delta}$	0.13	1.79	0.07	0.31
	WISC digit span backward	0.18	2.36	0.02	0.31

$^{\Delta}$ *indicates marginal significance level.*

Table 6 | Correlations between executive function scores and source accuracy in the serial and parallel interview conditions.

Interview condition	Inhibitory control tasks		Working memory tasks	
	Luria's hand game	Day/night stroop	WISC digit forward	WISC digit backward
Serial	0.04 (−0.01)	0.12 (−0.01)	0.31** (0.15)	0.31** (0.10)
Parallel	0.11 (0.08)	0.23** (0.14)	0.16 (0.05)	0.27** (0.15)

Partial correlations controlling for age are in parentheses.

**Significant at the 0.05 level (2-tailed).*

***Significant at the 0.01 level (2-tailed).*

For source accuracy scores, there were significant effects of working memory, but not of the inhibitory control variables. Working memory scores explained 9% of the variance in source accuracy scores. When age was entered into the regression, no significant predictors except for age remained.

The relationship between executive function and source accuracy as a function of retrieval strategy

The relationship between working memory and inhibitory control in the serial and parallel conditions was examined. To do so, separate correlations were run between the source scores of children in the serial and parallel conditions and the executive function measures, across age. All correlations for both conditions can be found in **Table 6**. For children in the serial condition, there were no correlations between the inhibitory control measures and source accuracy scores ($rs \leq 0.12$, $ps \geq 0.19$), but both working memory scores were significantly correlated with source accuracy. When these correlations were rerun as partial correlations controlling for age, only the WISC Backward scores were marginally related.

For children in the parallel condition, source accuracy scores were correlated with scores on the Stroop task and scores on the Digit Span Backward test, but not with Luria's Hand Game or Forward Digit Span scores. When age was controlled, these relationships were no longer significant ($p = 0.11$ and 0.10 for Stroop and Digit Span Backward, respectively).

DISCUSSION

The main purpose of this study was to evaluate the role of executive function as a predictor of episodic memory after exposure to multiple sources and of source-monitoring ability. In one condition, source monitoring was facilitated for the children by asking them to compare the two different sources ("real-life demonstration" and "story"); other children were asked about sources serially (i.e., they recalled details from one source first, and afterwards the other source). We expected that measures of working memory and inhibitory control would be related to both recognition and source accuracy, but that these relationships might be different in the serial and parallel conditions. Additionally, we expected that source decisions would be more accurate when details from the sources were recalled in parallel than when they were recalled serially.

DEVELOPMENTAL DIFFERENCES

We wanted to verify that, consistent with a large body of literature showing developmental changes between ages 4 and 8, there would be improvements in recognition, source monitoring and executive function; indeed we did find such patterns. While this is a replication of the majority of the findings in this area, it was important to establish that the pattern was the same in our particular source-monitoring tasks. Increases in accuracy with age provided the necessary data to test for relations between executive function and source monitoring. Importantly, all age groups performed well at identifying the details from the event, and all scored above chance on the source-monitoring task. Thus, the children in our sample were genuinely remembering the details and trying to identify their sources.

There was also an effect of source presentation. Source accuracy scores were better for the real-life demonstration than the story, indicating that this source was more salient for the children. In addition, children showed a "real-life bias"; that is, a bias toward reporting that details had come from the real-life demonstration more often. This is evidence of familiarity-based processing, because children reasoned that if they remembered seeing a detail, it must have happened in "real-life."

MEMORY, SOURCE MONITORING, AND EXECUTIVE FUNCTION

Two components of executive function were examined, and there was support for the hypothesis that both recognition and source monitoring are significantly correlated with measures of working memory and inhibitory control. Higher executive function scores were associated with better initial memory for the event, better delayed recognition of the details, and better identification of the sources. Regression analyses revealed that both working memory and inhibitory control were predictive of memory for event details immediately after the event and after a delay, but only working memory predicted source accuracy.

Generally, the significant relationships we found were weaker or non-significant when age was controlled for, which suggests that although executive function is related to recognition and source monitoring, a more general "cognitive development" factor is a stronger predictor than executive function alone. Clearly there are relationships between executive function, especially working memory, and source monitoring, as well as between executive function and recognition memory both immediately and after a delay. However, age as a construct represents improvements in many developmental processes, including theory of mind and reasoning about conflicting mental representations, which have also been shown to account for variance in source monitoring (Welch-Ross et al., 1997; Welch-Ross, 1999; Bright-Paul et al., 2008). Because age is tied to executive function as well as other cognitive abilities that are important for source monitoring, it is of course a stronger predictor than executive function alone.

The relationship between executive function and baseline accuracy scores suggests that executive function may play a role not only at retrieval, but also at encoding; those children with higher executive function scores recalled more accurate information about the event when there was no delay, and hence very little forgetting. Source monitoring may be enhanced by this initial processing because it could also be necessary at encoding to bind together features that allow source to be encoded with the memory, or enough suitable information to reason about source later. This is consistent with an argument from a recent review by Mammarella and Fairfield (2008) that working memory is important at encoding for binding the features of events together, which is crucial for source monitoring.

We hypothesized that working memory would be related to source-monitoring accuracy and this was supported. We reasoned that working memory may be necessary for holding information about different sources in mind, and engaging in "compare and contrast" reasoning. That is, if the two sources (live event and story) were recalled by the children, they would need to compare these sources with each other in order to decide the correct source. The results are similar to Ruffman et al.'s (2001) work, which showed that working memory was related to both recognition and source-monitoring accuracy. Ruffman et al. (2001) proposed that working memory plays a general role in memory ability that applies to recognition as well as source monitoring, rather than a differentiated effect on source monitoring alone.

It was hypothesized that children with better inhibitory control would be more accurate at source monitoring because they would be able to inhibit information from competing sources, and there was moderate support for this hypothesis. Although we did find that inhibitory control was positively correlated with source accuracy, interestingly, inhibitory control was not a significant predictor of source-monitoring accuracy in the regression analysis. The lack of variability in inhibitory control scores may have contributed to the non-significant findings in analyses with these variables. This issue is discussed further in the limitation section.

Similar to our results, Ruffman et al. (2001) and Melinder et al. (2006) found relationships between inhibitory control and some types of source-monitoring tasks, but not others. Ruffman et al. (2001) exposed children to audio and video stories, and showed a significant correlation between a Stroop task and source questions about details that happened in the video or in neither source, but no relationship to questions about details that happened in the audio or both sources. Melinder et al. (2006) found that while inhibitory control was a significant predictor of suggestibility, it was not predictive of source monitoring. Thus, while the inhibitory control-source monitoring relationship is theoretically plausible and evidence for the relationship is present in the literature, it is neither clear nor overwhelming.

THE CONTRIBUTIONS OF EXECUTIVE FUNCTION IN THE SERIAL AND PARALLEL CONDITIONS

One possible explanation for differences in results between various studies is variations in the way the source tasks are presented. We expected that there would be differences in the relationship between executive function and source monitoring as a function of retrieval strategy, and indeed there were differentiated relationships between components of executive function and source monitoring. Specifically, working memory was important for both tasks, but inhibitory control was only related to source monitoring with a parallel approach.

In the serial interview condition, children were asked to think about only one source at a time. Clearly working memory would be involved in remembering this rule, but we also predicted that inhibitory control would be important as children were required to inhibit competing information from the other source. As well, they would have to inhibit simple familiarity-based processes as they had to "filter" their memories, including details in their report only if a remembered detail was *also* accompanied by a determination of the target source, rather than anything that was at the activities.

Of relevance to this null result is the fact that source accuracy was lower in the serial condition compared to the parallel condition. Thus, it is possible that children in the serial condition were simply remembering information with less regard to source than their counterparts (i.e., failing to "filter" through source). While all age groups scored above chance in source monitoring, it is possible that children in the serial condition were simply engaging in less source reasoning than those in the parallel condition. Therefore, an inhibitory control-source monitoring relationship would be less apparent in the serial condition if children were not engaging as extensively in source monitoring processes.

Scores on the Stroop task were related to source monitoring accuracy in the parallel condition. This is consistent with a previous study showing that inhibitory control was related to resistance to suggestions about a series of repeated events (Roberts and Powell, 2005). Although we had originally anticipated that inhibitory control would play a stronger role in the serial condition, where children were required to inhibit details from other sources, it is clear that inhibition serves a useful function when children are making decisions about several competing sources as well. In our parallel processing task, children sorted cards between several different boxes, and the presentation of competing source options may have required inhibitory control as well as working memory.

Perhaps the fact that young children are not proficient in inhibitory control may underlie their lack of spontaneously recalling sources in parallel. This is supported by the finding that the younger age groups (4- to 5 and 6-year-olds) who were provided with a "compare and contrast" strategy improved their source monitoring relative to those practicing a serial retrieval strategy. The relationship might also be bidirectional so that engaging in a parallel retrieval strategy necessitates an improvement in inhibitory control.

Young children can monitor self-other sources before they can monitor two internally generated events (e.g., imagined and dreamt events; Foley et al., 1983; Foley and Johnson, 1985). Similarly, young children are disproportionately less able to monitor sources that are similar compared to older children and adults (Lindsay et al., 1991; Roberts and Blades, 1999). These findings demonstrate clearly that the demands of source-monitoring tasks have diverse influences on accuracy resulting in different developmental patterns. Thus, it may be fruitful to consider what factors contribute to task difficulty and interweave this with investigations of inhibitory control-source monitoring relationships. Careful study of the characteristics of source tasks and how they influence the role of the cognitive factors involved in source monitoring is a necessary step to better understanding the executive underpinnings of source-monitoring development.

SERIAL vs. PARALLEL SOURCE ACCURACY

We hypothesized that there would be differences in source accuracy when children recalled sources serially vs. in parallel, with the parallel condition showing an advantage over the serial condition. This was true for the two younger age groups but not for the older children. Accuracy scores were very similar across conditions for the 7- to 8-year-olds, and there was a large difference between the scores for 4- and 5-year-olds in the two conditions; the children in the serial condition demonstrated poor source-monitoring abilities, and the children in the parallel condition improved by 15%, bringing their performance close to that of the 7- to 8-year-olds.

When young children considered both sources at the same time during the decision-making process, they monitored source more carefully and benefitted from the facilitation of a comparison strategy. In the serial condition children were provided with the opportunity to spontaneously use a strategy, but were not assisted with comparing sources. We believe that differences between serial and parallel retrieval strategies were not evident for the 7- to 8-year-old group because these children were able to spontaneously engage in parallel retrieval of sources without the interviewer facilitating such a strategy. Developmentally, it is around this time that children are close to adult proficiency in some types of source recall (Roberts, 2002).

PRACTICAL IMPLICATIONS

The results of this study have implications for educational and forensic contexts. Younger children may be preoccupied with absorbing content rather than source information because it is more important for young children to build up a knowledge base, and only in later years concern themselves with recalling where information came from (Roberts and Powell, 2005). This lack of attention to source has been well documented in several different areas of cognitive development (e.g., Gopnik and Graf, 1988). In contrast, older children who have built up a significant (though by no means complete) knowledge of the world have more cognitive resources available for attending to the sources of information. Indeed, as children become habitual internet-users and the availability of information grows, making judgments about the credibility of information from different sources will serve children well.

These findings are also relevant to forensic investigations involving children. For example, many children in abuse investigations are asked to provide specific information about an alleged incident, which requires them to distinguish between instances because child abuse often occurs more than once (Ceci and Bruck, 1993). Children might confuse details from similar events that happened a long time ago because they confuse the origins of events (Roberts and Blades, 1999). Developing techniques that compensate for young children's still-developing proficiency in executive function and source monitoring is a difficult but especially important challenge.

Investigators may be able to encourage children to directly compare sources and think carefully about multiple instances before deciding in which event a detail occurred in order to increase source accuracy. Most children with experiences of repeated abuse will have built up a script and may not realize the importance of reporting details specific to just one instance, so drawing children's attention to sources in this way may facilitate source monitoring performance. For example, Brubacher et al. (2011) have found that giving children practice in talking about occurrences of a repeated event (e.g., swimming lessons) improved their reports when asked to discuss target instances of another repeated event. The fact that the parallel retrieval strategy in this study improved source monitoring through the task procedure alone without a separate training procedure makes this technique ideal for investigators, as it requires few resources to employ. However, more research on the effectiveness of this technique is needed before generalizations are made.

LIMITATIONS

A limitation of the current study was that the inhibitory control scores showed evidence of ceiling effects. Inhibitory control shows rapid improvements in early childhood, with the largest improvements in tasks like Luria's Hand Game around age 4 (Best and Miller, 2010). Therefore, it would be expected that 6- to 8-year-olds would have similarly high scores on these tasks, whereas the 4- to 5-year-olds would not have scores as high as the older children. Although the scores in inhibitory control tasks were quite high in this study, the relationships with age and working memory were significant, so it was not the case that variability was so restricted that it was not possible to find significant relationships with other variables. Reaction time data were not available in our study, but this type of data might be considered more useful for future research as it may show more variability and be less susceptible to ceiling effects.

Another limitation of this study is that several aspects of the methodology may have reduced the demands of executive function in the current source-monitoring tasks. The boxes labeled with the source names may have reduced the need for working

memory because children were not required to hold the possible source options in mind as they thought about details in the way they would have been if the questions were asked verbally. As well, our two-step memory task may have reduced demands on executive function compared to a task where recognition and source were combined using "Story," "Real-Life," and "Neither" options, because in these tasks children are required to think about whether they saw a detail and what source it was in at the same time. However, although this two-step procedure may be less cognitively demanding overall, it allowed for an investigation of the relationships of executive function to recognition and source-monitoring accuracy separately.

CONCLUSION

This study adds evidence to the growing body of literature on the underlying mechanisms of source-monitoring development and, overall, these findings have illustrated the relations between executive function and source-monitoring accuracy. Working memory seems to be necessary for source monitoring in general, even when the exact nature of the task varies. The role of inhibitory control in source monitoring is less clear, although inhibitory control was positively correlated with memory and source accuracy. Further research is necessary to clarify mixed results about the contributions of working memory and inhibitory control to source-monitoring performance in previous research. Although this study contributes to the body of literature on this topic, it does not ultimately provide a definitive answer to that question.

The results of this research address both practical and theoretical questions about what interview strategies are most helpful for children when they are making source-monitoring decisions. Knowing more about the cognitive prerequisites for source monitoring helps determine what to expect from children of different ages and cognitive abilities. An important area for future research is the investigation of how task difficulty affects the relationship between executive function and source monitoring.

AUTHOR CONTRIBUTIONS

Both authors contributed ideas to the development of the project. Data was collected, coded and analyzed by Becky Earhart in collaboration with a team of research assistants. Both authors made significant contributions to the writing of this manuscript.

ACKNOWLEDGMENTS

This research was conducted for the Masters Thesis of Becky Earhart. The research was funded by a Natural Sciences and Engineering Research Council (NSERC) Grant (#249862) to Kim P. Roberts. Portions of this research were presented to the 2013 annual meeting of the American Psychology-Law Society and the 2013 biennial meeting of the Society for Research in Child Development. We are grateful to the Waterloo Region District School Board and the families in the Waterloo region who participated in this research, as well as the research assistants who contributed to the project: Sonja Brubacher, Donna Drohan-Jennings, Courtney Arseneau, Katherine Wood, Kayleen Willemsen, Sam Chefero, Candice Sommers, Brittney Dudar, and Paula Ghelman.

REFERENCES

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Bright-Paul, A., Jarrold, C., and Wright, D. B. (2005). Age-appropriate cues facilitate source monitoring and reduce suggestibility in 3- to 7-year-olds. *Cogn. Dev.* 20, 1–18. doi: 10.1016/j.cogdev.2004.06.001

Bright-Paul, A., Jarrold, C., and Wright, D. B. (2008). Theory-of-mind development influences suggestibility and source monitoring. *Dev. Psychol.* 44, 1055–1068. doi: 10.1037/0012-1649.44.4.1055

Brubacher, S. P., Roberts, K. P., and Powell, M. B. (2011). Effects of practicing episodic versus scripted recall on children's subsequent narratives of a repeated event. *Psychol. Public Policy Law* 17, 286–314. doi: 10.1037/a0022793

Bruck, M., and Melnyk, L. (2004). Individual differences in children's suggestibility: a review and synthesis. *Appl. Cogn. Psychol.* 18, 947–996. doi: 10.1002/acp.1070

Ceci, S. J., and Bruck, M. (1993). Suggestibility of the child witness: a historical review and synthesis. *Psychol. Bull.* 113, 403–439. doi: 10.1037/00332909.113.3.403

De Luca, C. R., and Leventer, R. J. (2008). *Developmental Trajectories of Executive Functions Across the Lifespan*. Philadelphia, PA: Taylor & Francis.

Fahie, C. M., and Symons, D. K. (2003). Executive functioning and theory of mind in children clinically referred for attention and behavior problems. *J. Appl. Dev. Psychol.* 24, 51–73. doi: 10.1016/S01933973(03)00024-8

Foley, M. A., and Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: a developmental comparison. *Child Dev.* 56, 1145–1155. doi: 10.2307/1130229

Foley, M. A., Johnson, M. K., and Raye, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Dev.* 54, 51–60. doi: 10.2307/1129860

Gerrie, M. P., and Garry, M. (2007). Individual differences in working memory capacity affect false memories for missing aspects of events. *Memory* 15, 561–571. doi: 10.1080/09658210701391634

Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action: performance of children 3 1/2-7 years old on a stroop-like day-night test. *Cognition* 53, 129–153.

Glisky, E. L., Polster, M. R., and Routhieaux, B. C. (1995). Double dissociation between item and source memory. *Neuropsychology* 9, 229–235. doi: 10.1037/0894-4105.9.2.229

Gopnik, A., and Graf, P. (1988). Knowing how you know: young children's ability to identify and remember the sources of their beliefs. *Child Dev.* 59, 1366–1371. doi: 10.2307/ 1130499

Hughes, C. (1996). Control of action and thought: normal development and dysfunction in autism: a research note. *J. Child Psychol. Psychiatry* 37, 229–236. doi: 10.1111/j.1469-7610.1996.tb01396.x

Jaschinski, U., and Wentura, D. (2002). Misleading postevent information and working memory capacity: an individual differences approach to eyewitness memory. *Appl. Cogn. Psychol.* 16, 223–231. doi: 10.1002/acp.783

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3

Kanakogi, Y., Moriguchi, Y., Fu, G., Lee, K., and Itakura, S. (2012). How does executive function contribute to source monitoring in young children? *Psychologia* 55, 194–207. doi: 10.2117/psysoc.2012.194

Karpinski, A. C., and Scullin, M. H. (2009). Suggestibility under pressure: theory of mind, executive function, and suggestibility in preschoolers. *J. Appl. Dev. Psychol.* 30, 749–763. doi: 10.1016/j.appdev.2009.05.004

Lindsay, S., Johnson, M. K., and Kwon, P. (1991). Developmental changes in memory source monitoring. *J. Exp. Child Psychol.* 52, 297–318. doi: 10.1016/0022-0965(91)90065-Z

Mammarella, N., and Fairfield, B. (2008). Source monitoring: the importance of feature binding at encoding. *Eur. J. Cogn. Psychol.* 20, 267–273. doi: 10.1080/09541440601112522

Melinder, A., Endestad, T., and Magnussen, S. (2006). Relations between episodic memory, suggestibility, theory of mind, and cognitive inhibition in the preschool child. *Scand. J. Psychol.* 47, 485–495. doi: 10.1111/j.1467-9450.2006.00542.x

Powell, M. B., and Thomson, D. M. (1996). Children's memory of an occurrence of a repeated event: effects of age, repetition, and retention interval across three question types. *Child Dev.* 67, 1988–2004. doi: 10.2307/1131605

Powell, M. B., and Thomson, D. M. (2003). Improving children's recall of an occurrence of a repeated event: is it a matter of helping them to generate options? *Law Hum. Behav.* 27, 365–384. doi: 10.1023/A:1024032932556

Reck, S. G., and Hund, A. M. (2011). Sustained attention and age predict inhibitory control during early childhood. *J. Exp. Child Psychol.* 108, 504–512. doi: 10.1016/j.jecp.2010.07.010

Roberts, K. P. (2002). Children's ability to distinguish between memories from multiple sources: implications for the quality and accuracy of eyewitness statements. *Dev. Rev.* 22, 403–435. doi: 10.1016/S0273-2297(02)00005-9

Roberts, K. P., and Blades, M. (1999). Children's memory and source monitoring of real-life and televised events. *J. Appl. Dev. Psychol.* 20, 575–596. doi: 10.1016/S0193-3973(99)00030-1

Roberts, K. P., and Powell, M. B. (2005). The relation between inhibitory control and children's eyewitness memory. *Appl. Cogn. Psychol.* 19, 1003–1018. doi: 10.1002/acp.1141

Ruffman, T., Rustin, C., Garnham, W., and Parkin, A. J. (2001). Source monitoring and false memories in children: relation to certainty and executive functioning. *J. Exp. Child Psychol.* 80, 95–111. doi: 10.1006/jecp.2001.2632

Schacter, D. L., Kagan, J., and Leichtman, M. D. (1995). True and false memories in children and adults: a cognitive neuroscience perspective. *Psychol. Public Policy Law* 1, 411–428. doi: 10.1037/1076-8971.1.2.411

Thierry, K. L., Goh, C. L., Pipe, M. E., and Murray, J. (2005). Source recall enhances children's discrimination of seen and head events. *J. Exp. Psychol. Appl.* 11, 33–44. doi: 10.1037/1076-898X.11.1.33

Thierry, K. L., Spence, M. J., and Memon, A. (2001). Before misinformation is encountered: source monitoring decreases child witness suggestibility. *J. Cogn. Dev.* 2, 1–26. doi: 10.1207/S15327647JCD0201_1

Unsworth, N., and Brewer, G. A. (2010). Variation in working memory capacity and intrusions: differences in generation or editing? *Eur. J. Cogn. Psychol.* 22, 990–1000. doi: 10.1080/09541440903175086

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition.* San Antonio, TX: The Psychological Corporation.

Welch-Ross, M. K. (1999). Interviewer knowledge and preschoolers' reasoning about knowledge states moderate suggestibility. *Cogn. Dev.* 14, 423–442. doi: 10.1016/S0885-2014(99)00013-1

Welch-Ross, M. K., Diecidue, K., and Miller, S. A. (1997). Young children's understanding of conflicting mental representation predicts suggestibility. *Dev. Psychol.* 33, 43–53. doi: 10.1037/0012-1649.33.1.43

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 January 2014; accepted: 17 April 2014; published online: 08 May 2014.
Citation: Earhart B and Roberts KP (2014) The role of executive function in children's source monitoring with varying retrieval strategies. Front. Psychol. ***5****:405. doi: 10.3389/fpsyg.2014.00405*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Earhart and Roberts. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 18 June 2014
doi: 10.3389/fpsyg.2014.00616

Measuring inhibitory control in children and adults: brain imaging and mental chronometry

***Olivier Houdé*[1,2] * *and Grégoire Borst*[1]**

[1] CNRS Unit 8140, Laboratory for the Psychology of Child Development and Education, Alliance for Higher Education and Research Sorbonne-Paris-Cité, Paris Descartes University, Paris, France
[2] Institut Universitaire de France, Paris, France

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Andrew Simpson, University of Essex, UK
Melanie Stollstorff, University of Colorado Boulder, USA

***Correspondence:**
Olivier Houdé, CNRS Unit 8140, Laboratory for the Psychology of Child Development and Education, Alliance for Higher Education and Research Sorbonne-Paris-Cité, Paris Descartes University, 46 rue Saint-Jacques, 75005 Paris, France
e-mail: olivier.houde@parisdescartes.fr

Jean Piaget underestimated the cognitive capabilities of infants, preschoolers, and elementary schoolchildren, and overestimated the capabilities of adolescents and even adults which are often biased by illogical intuitions and overlearned strategies (i.e., "fast thinking" in Daniel Kahneman's words). The crucial question is now to understand why, despite rich precocious knowledge about physical and mathematical principles observed over the last three decades in infants and young children, older children, adolescents and even adults are nevertheless so often bad reasoners. We propose that inhibition of less sophisticated solutions (or heuristics) by the prefrontal cortex is a domain-general executive ability that supports children's conceptual insights associated with more advanced Piagetian stages, such as number-conservation and class inclusion. Moreover, this executive ability remains critical throughout the whole life and even adults may sometimes need "prefrontal pedagogy" in order to learn inhibiting intuitive heuristics (or biases) in deductive reasoning tasks. Here we highlight some of the discoveries from our lab in the field of cognitive development relying on two methodologies used for measuring inhibitory control: brain imaging and mental chronometry (i.e., the negative priming paradigm). We also show that this new approach opens an avenue for re-examining persistent errors in standard classroom-learning tasks.

Keywords: inhibition, conceptual development, brain imaging, negative priming, number, categorization, logical reasoning

The scientific study of cognitive development in young children traces its roots back to Jean Piaget, a pioneer of this field in the 20th century (Piaget, 1954, 1983). Piaget described children as active learners who, through numerous interactions with their environments, construct a complex understanding of the physical world around them. From infancy to adolescence, children progress through four psychological stages: (1) the sensorimotor stage from birth to 2 years (when cognitive functioning is based primarily on biological reactions, motor skills and perceptions); (2) the preoperational stage from 2 to 7 years (when symbolic thought and language become prevalent, but reasoning is illogical by adult standards); (3) the concrete operations stage from 7 to 12 years (when logical reasoning abilities emerge but are limited to concrete objects and events); and (4) the formal operations stage at ∼12 years (when thinking about abstract, hypothetical, and contrary-to-fact ideas becomes possible).

FROM PIAGET'S THEORY TO INHIBITORY CONTROL MODEL

Piaget underestimated the cognitive capabilities of infants, preschoolers, and elementary schoolchildren, and he overestimated the capabilities of adolescents and adults, which are often biased by illogical intuitions and overlearned strategies (or heuristics) they fail to inhibit (Houdé, 2000, 2014; Kahneman, 2011). During the last three decades, detailed behavioral studies of children's problem solving led to a reconceptualization of cognitive development, from discrete Piagetian stages to one that is analogous to overlapping waves (Siegler, 1996, 1999). The latter is consistent with a neo-Piagetian approach of cognitive development, in which more and less sophisticated solutions compete for expression in the human brain. In this approach, inhibition of less sophisticated solutions by the prefrontal cortex is a critical component of children's conceptual insights associated with more advanced Piagetian stages (Houdé et al., 2000, 2011; Poirel et al., 2012; Borst et al., 2013a). According to this theoretical framework, the development of inhibitory control efficiency during childhood and adolescence contributes to the development of conceptual knowledge in various cognitive domains. This view is consistent with a number of studies showing that the dramatic development of the inhibitory control efficiency between 3- and 5-years old (e.g., Carlson, 2005) explains to some extent the growing ability of children to succeed in Theory of Mind (e.g., Benson et al., 2013), counterfactual reasoning (e.g., Beck et al., 2009) and strategic reasoning (e.g., Apperly and Carroll, 2009) tasks. In both of these literatures inhibition is viewed as a domain-general process allowing children and adults to resist habits or automatisms, temptations, distractions, or interference, and to adapt to conflicting situations (Diamond, 2013). Finally, in our view the gradual improvement of cognitive abilities in different domains is directly related to the improvement of inhibitory control efficiency. Note, however, that the development

of inhibitory control efficiency is necessary but probably not sufficient to produce conceptual development during childhood and adolescence.

At any point in time, children and adults potentially have available to them heuristics (i.e., intuitions) and logicomathematical algorithms, or as Kahneman (2011) described, multiple levels of "thinking fast and slow." Heuristics are rapid, often global or holistic, useful strategies in many situations, *but sometimes they are misleading*, whereas algorithms are slow, demanding and analytical strategies that *necessarily lead to a correct* (i.e., *logical*) *solution in every situation*. In general, children and adults prefer using fast heuristics spontaneously, but that choice does not indicate that they are illogical *per se* (Houdé, 2000) or that they are "happy fools" (De Neys et al., 2013, 2014). Psychologists had to be careful to avoid false negatives (Gelman, 1997), which is a strong tendency to say that those children or adults who fail a task are incompetent in the target domain of knowledge. A "presumption of rationality" is sometimes the best assessment.

Contrary to Piaget's theory, infants learn more about the outside world through information that is captured by their perceptual systems than through motor skills development (Mandler, 1988; Baillargeon, 1995; Spelke, 2000). Infant cognition studies evaluated both the capacity to interpret sensory data and the faculty for understanding and reasoning about complex events. In the last 10 years, theoretical ideas and empirical research in the field have demonstrated that very young children's learning and thinking mechanisms do remarkably resemble the basic inductive processes of science, i.e., probabilistic models and Bayesian learning methods (Gopnik, 2012). Infants can implicitly reason statistically (Téglas et al., 2011). From this point of view, the very young child is already seen as a "scientist in the crib" (Gopnik et al., 1999).

THE NUMBER EXAMPLE

A heated debate topic in psychology is how children come to understand numbers (Dehaene, 1997). Piaget's answer was that number is constructed in children through the logicomathematical synthesis of classification and seriation operations (Piaget, 1952). Number borrows its inclusion structure from classes (1 is included in 2, 2 in 3, etc.); because it disregards qualities by transforming objects into units, it brings a serial order into play, the sole means of distinguishing one unit from the next: 1 then 1, then 1, etc. The serial ordering of units is combined with the inclusion of the sets that result from their union (1 is included in $1 + 1$, $1 + 1$ is included in $1 + 1 + 1$, etc.) to constitute number. The task Piaget used was conservation of number. When children are shown two rows of objects that contain an equal number of objects but that differ in length (because the objects in one of the rows have been spread apart), young children think the longer row has more objects. Piaget's interpretation was that preschool children are still fundamentally intuitive (their reasoning is illogical by adult standards), or as he called them, "preoperational" (Stage 2), and hence limited to a perceptual way of processing information (based on length or, in certain cases, on density). At the age of 6 or 7 years old, children understand the equivalency of quantities, regardless of apparent transformations. At this point, they are called "operational" or "conserving" (Stage 3), the criterion for mastery of number. Piaget also worked on determining whether the conservation of number develops simultaneously with inclusion (classification) and order relations (seriation).

After this founding work on the genesis of number, research in this domain proliferated, and criticisms of Piaget's theory were far from scarce. First, the synchronous development of classification, seriation, and conservation was not validated in experimental verifications. Second, it became clear that Piaget's view of the logico-structural aspect of number is overly polarized and overshadows the more functional aspects of numerical development, such as counting.

A radical change in perspective began with Gelman and Meck (1983), Gelman et al. (1986), who not only turned the attention toward counting but also postulated the early existence of five fundamental principles of counting: stable order (order of the number words), strict one-to-one correspondence (between the number words and the items counted), cardinality (the number word corresponding to the last item counted is equal to the total number of items), abstraction (any kind of item can be counted), and order irrelevance (items can be counted in any order). Gelman demonstrated the presence of these principles in young children by having them say whether they thought a doll was counting correctly or incorrectly. Knowledge or lack of knowledge of a given principle was deduced from whether the child detected the corresponding type of counting error (unstable order, violation of the one-to-one correspondence, cardinal number referred to by an ordinal word number, etc.). The results indicated that 3-year-old children have already acquired the basic principles of counting. This led Gelman to distinguish three components in the ability to count: a conceptual component ("knowing why" or understanding the five principles), a procedural component ("knowing how" or understanding the structure and order of counting), and a utilization component ("knowing when" or understanding the relevance of using the first two components in a given context). Defending the principles-before-skills hypothesis, Gelman suggested that the numerical difficulties of preschool children lie essentially in the procedural and utilization components. Another of Gelman's original contributions was her use of the so-called "magic task" to demonstrate that 3- to 4-year-old children are surprised by transformations that affect the cardinal number of a set (adding and subtracting items) but not by transformations that do not (spreading and grouping) (Gelman, 1972). She concluded that despite their failure in Piaget's conservation of number task, the children at this age are already capable of seeing through irrelevant transformations and treating the number of items as invariable (for a seminal study on this point, see Mehler and Bever, 1967). This new conclusion was corroborated by the discovery of the perception of numerical invariance in neonates (Antell and Keating, 1983) and in 5- and 8-month-old infants (Loosbroek and Smitsman, 1990; Lipton and Spelke, 2003).

The most striking example of infants as "mathematicians" is found in the famous work by Wynn (1992). Wynn recorded the looking time of 4- and 5-month-old infants in the "impossible-event" procedure (or violation-of-expectation procedure) and demonstrated that infants were surprised by (looked longer at) impossible numerical events (e.g., $1 + 1 = 1$ and $1 + 1 = 3$,

or 2 − 1 = 2) but were not surprised at the corresponding possible events (1 + 1 = 2 and 2 − 1 = 1; the events were staged with Mickey Mouse figures). She concluded that infants are endowed with a mechanism that calculates the exact outcome of simple arithmetic operations. She claimed that infants at this age are already able to encode ordinal information and possess genuine numerical concepts that cannot be reduced to holistic percepts derived from a pattern recognition process (for a brain imaging confirmation of Wynn's results, see Berger et al., 2006). Like Gelman's stance, Wynn's position is strong and seems to run counter to what we know about the numerical difficulties of preschool children. Wynn's empirical results are robust and consistent (Wynn, 2000), but they have sparked theoretical debates (Simon, 1997, 1998). The task of the following research was to devise a developmental model of logicomathematical operations (conservation, counting, and elementary arithmetic) that accounts for both early abilities (Gelman, Wynn, etc.) and late inabilities (Piaget), without denying the reality of the former but raising the question of the factors that explain the latter.

HEURISTICS, ALGORITHMS, COGNITIVE CONTROL, AND INHIBITION OF MISLEADING STRATEGIES

How might we explain the famous number-conservation error observed in children until the age of seven by Piaget and, after him, by all developmental psychologists around the world? It is an intriguing question because we know today that very young children are already capable of treating the number of items as invariable through irrelevant transformations and that they possess other protonumerical skills. One of the main current explanations is that children learn heuristics, which are often useful in a large set of situations, but fail to inhibit them when, contrary to general practice, they are misleading (Houdé, 2000, 2014). In the case of Piaget's number-conservation algorithm, the overlearned competing heuristic is "length-equals-number" (Houdé and Guichart, 2001). This new theoretical approach is in line with Diamond's explanation of the A-not-B error in infants (see Diamond, 1991) and assumes that cognitive development relies not only on the acquisition of knowledge of incremental complexity (Piaget, 1983) but also on the ability to inhibit previously acquired knowledge (Bjorklund and Harnishfeger, 1990; Diamond, 1991, 1998; Dempster and Brainerd, 1995; Harnishfeger, 1995; Houdé, 2000). Increasing evidence shows that the ability to inhibit previous knowledge is critical for developmental milestones, such as those defined by Piaget's theory (Borst et al., 2013a; Houdé, 2014). Inhibitory control of misleading strategies, an executive function performed by the prefrontal cortex, has been claimed necessary for acquisition and use of motor or cognitive algorithms in the fields of object permanence in infants (Diamond and Goldman-Rakic, 1989; Diamond, 1991, 1998; Bell and Fox, 1992), number-conservation and class inclusion in preschool and schoolchildren (Houdé and Guichart, 2001; Perret et al., 2003; Borst et al., 2012, 2013b), and logical reasoning in adolescents and adults (Houdé et al., 2000; Houdé and Tzourio-Mazoyer, 2003; Houdé, 2007).

One of the challenges of today's developmental research, in all domains of cognition ranging from motor programming to high-order logical reasoning, is to account not only for a general and incremental process of coordination-activation capacities of structural units, schemes or skills through ages and stages (Piaget, 1983, and all the 1980s neo-Piagetians: see the review book by Demetriou, 1988) but also for a general process of selection-inhibition of competing strategies, i.e., heuristics (or intuitions) and logicomathematical algorithms, occurring with different weights at any point in time, depending on the context, in a non-linear dynamical system of growth (Siegler, 1996, 1999; Houdé, 2000, 2014). Such cognitive model introduces less regular developmental curves containing perturbations, bursts, and collapses. O'Reilly (1998) described six principles for biologically based computational models of cognition, one of which is inhibitory competition (see also Johnson, 2010). Resolving this "inhibition issue" is an important task for both developmental psychology and cognitive neuroscience. The most compelling magnetic resonance imaging (MRI) reports of structural changes with brain development during childhood and adolescence showed a sequence in which the higher-order association area, such as the prefrontal cortex sustaining inhibitory control, matures last (Casey et al., 2005). The sequence in which the cortex matures parallels the cognitive milestones in human development. First, the regions subserving primary functions, such as motor and sensory systems, mature the earliest; the temporal and parietal association cortices associated with basic language skills and spatial attention mature next; and the last to mature are the prefrontal cortex and its inhibitory control ability.

BRAIN IMAGING: INHIBITORY CONTROL AND PREFRONTAL CORTEX

Using fMRI (functional magnetic resonance imaging), from this theoretical perspective, we re-examined what occurs in the developing brain when school children are tested for their performance in Piaget's number-conservation task. Remember that when children are shown two rows of objects that contain an equal number of objects but that differs in length (because the objects in one of the rows had been spread apart), young children think that the longer one has more objects. Piaget's interpretation was that preschool children are still fundamentally intuitive (their reasoning being illogical by adult standards), or as he called them, "preoperational" (Stage 2), and hence limited to a perceptual way of processing information (here, based on length or, in certain cases, on density). When they are ∼6 or 7 years old, children understand the equivalency of quantities, regardless of apparent transformations. At this point, they are called "operational" or "conserving" (Stage 3), the criterion for logicomathematical mastery of number. Our new hypothesis was that their main cognitive difficulty (beyond logicomathematical cognition *per se*) was to efficiently inhibit through their prefrontal cortex the overlearned "length-equals-number" strategy, a heuristic that is often used both by children and adults in many school and everyday situations.

In a first fMRI study, we found that the cognitive change allowing children to access conservation (i.e., the shift from Stage 2 to Stage 3 in Piaget's theory) was related to the neural contribution of a bilateral parietofrontal network involved in

numerical and executive functions (Houdé et al., 2011). These imaging results highlighted how the behavioral and cognitive stages that Piaget formulated during the 20th century manifest in the brain with age. In a second fMRI study (Poirel et al., 2012), we demonstrated that the prefrontal activation (i.e., the blood-oxygen-level-dependent signal) observed when schoolchildren succeeded at the Piaget's number-conservation task was correlated to their behavioral performance on a Stroop-like measure of inhibitory function development (Wright et al., 2003). These new results in schoolchildren fit well with previous brain imaging data from our laboratory showing a key role of prefrontal inhibitory control training when adolescents or adults (belonging to Stage 4 in Piaget's theory) spontaneously fail to block their perceptual intuitions (or bias, heuristics) to activate logicomathematical algorithms (i.e., deductive rules) in reasoning tasks (Houdé et al., 2000; Houdé, 2007).

If we have "two minds in one brain" as stated by Evans (2003) or, in other words, two ways of thinking and reasoning, i.e., "fast and slow" (Kahneman, 2011), currently called "System 1" (intuitive system) and "System 2" (analytic system), then the crucial challenge is to learn to inhibit the misleading heuristics from System 1 when the more analytic and effortful System 2 (logicomathematical algorithms) is the way to solve the problem (Houdé, 2000, 2014; Borst et al., 2013a). Within this post-Piagetian theoretical approach, we can now understand why, despite rich precocious knowledge about physical and mathematical principles observed in infants and young children, older children, adolescents, and adults so often have poor reasoning. The cost of blocking our intuitions is high and depends on the late maturation of the prefrontal cortex. Moreover, this executive ability remains delicate throughout our lifetime, and adults may sometimes need "prefrontal pedagogy" to learn inhibiting intuitive heuristics (or biases) in reasoning tasks (Houdé, 2007).

An innovative research question now is to better understand the cognitive roots of such powerful heuristics (intuitions and bias from System 1) that children and adults have so much difficulty inhibiting in some cases. New heuristics may appear and be overlearned at any time in the course of development (Houdé, 2000, 2014) because our brain is an irrepressible detector of regularities from its perceptual and cultural environment. For example, preschool children (more than infants) are often exposed, in "math books" in the classroom or in everyday scenes, to patterns of objects in which number and length covary (e.g., the 1-to-10 Arabic numbering series is frequently illustrated by increasing lines of drawn animals or fruits: one giraffe, two hippopotamus, three crocodiles, and so on), hence the overlearned and misleading "length-equals-number" heuristic, which is overactivated in Piaget's conservation of number task. A new avenue of research would be to assess the role of early sensitivity to statistical patterns (i.e., probability of hypotheses) and Bayesian inference (Gopnik, 2012) in the psychological construction of perceptual, motor, and cognitive heuristics. Moreover, the power of Bayesian learning might require, in some conflict situations, a strong antagonist process of inhibition for blocking heuristics when they are misleading.

MENTAL CHRONOMETRY: INHIBITORY CONTROL AND NEGATIVE PRIMING EFFECT

In this section, we will review mental chronometry studies that used negative priming to demonstrate the role of inhibitory control in logicomathematical tasks. The logic of the negative priming approach is as follows: if information (or a perceptual or cognitive heuristic) was previously ignored (or inhibited), then the subsequent processing of that information (or the subsequent activation of that heuristic strategy) will be disrupted as revealed by slower or less accurate responses (see, e.g., Tipper, 1985, 2001; Neill et al., 1995). In the classical negative priming paradigm, participants performed pairs of stimuli. The first stimulus of the pair is the prime; the second one is the probe. Classically, participants' performance is measured on the second stimulus (i.e., probe). Critically, performance are compared between test-probes in which the target is a distractor inhibited on the first stimulus (i.e., prime) and control-probes in which the target bears no relation with a distractor inhibited on the prime. The logic of the negative approach is similar for strategies: if to reason logically one need to inhibit an overlearned strategy (or heuristic) to activate a logical algorithm, then a negative priming effect should be observed when participants perform prime-probe sequences in which the heuristic that needs to be activated on the probe was inhibited on the prime. Bluntly put, if people block the heuristic response on one trial, they will pay a price if they need to rely on it on the subsequent trial.

Following this logic, Houdé and Guichart (2001) devised the first negative priming paradigm to demonstrate that inhibitory control was required when children correctly solved a classic logicomathematical task – Piaget's number-conservation task (Piaget, 1952). The authors asked children to perform two types of prime-probe trials. In test trials, two rows of different length but with the same number of objects (i.e., a classical number-conservation item) were presented as the prime. In order to correctly state that the two rows contained the same number of objects, children had to inhibit the length-equals-number heuristic. On the probe, an item in which length and number co-varied – i.e., the longer row contained more objects – was displayed. Critically, the length-equals-number strategy that was inhibited on the prime became the appropriate strategy to activate on the probe. In control trials, the strategy to be used on the prime was unrelated to the strategy to activate on the probe. Objects were displayed in such a way that counting each object was the only appropriate strategy (i.e., the objects on one of the rows were displayed vertically on the screen which ruled out using the length-equals-number strategy). As on the test trials, an item in which length and number co-varied was displayed on the probe. Comparison of the probe response times between test and control trials revealed a clear negative priming effect: children were slower to use the length-equals-number strategy after they performed a typical Piaget-like number-conservation item in which the length-equals-number heuristic needs to be inhibited to overcome the interference between the length of the rows and the number of objects. This result suggests that children's ability to reason correctly on number-conservation tasks is directly related to their ability to inhibit a misleading strategy.

Note that as opposed to Piaget's seminal number-conservation task, the transformation (i.e., the lengthening of one of the rows) is not presented to the children in Houdé and Guichart (2001)' study (due to the inherent structure of such sequential paradigm). Thus, one could claim that Piaget's seminal number-conservation task and the Piaget-like number-conservation task designed by Houdé and Guichart test very different numerical knowledge of the children and (b) that success at Piaget's seminal number-conservation task might have nothing to do with the inhibition of the length-equals-number strategy. However, recent fMRI and high-density EEG studies (Houdé et al., 2011; Poirel et al., 2012; Borst et al., 2013c) revealed that children and adults must inhibit the length-equals-strategy to succeed at Piaget's seminal number-conservation task in agreement with the results reported on the Piaget-like number-conservation task designed by Houdé and Guichart (2001).

In a follow-up electrophysiological study using a similar negative priming adaptation of the number-conservation task with young adults, Daurignac et al. (2006) reported enhanced amplitude of the N200 wave (with a large distribution over the scalp) when the length-equals-number strategy inhibited on the prime became the appropriate strategy to activate on the probe. Given that the N200 is assumed to reflect inhibitory control, electrophysiological data garnered in this study suggest that adults as children need to inhibit the length-equals-number heuristic to reason correctly here (see Borst et al., 2013c for an incremental demonstration using high-density ERP).

Negative priming has also been reported in another famous logicomathematical Piagetian task, the class-inclusion task (Inhelder and Piaget, 1964). In this task, ten daisies (i.e., the subordinate class A) and two roses (i.e., the subordinate class A') are presented to the child and he(she) is asked whether there are more daisies than flowers (i.e., the superordinate class B = A+A'). Before the age of seven, children erroneously think that there are more daisies than flowers because they fail to perform the appropriate comparison between the superordinate class (flowers) and the subordinate class (daisies). To succeed at this task, children need to inhibit the direct (heuristic) perceptual comparison of the visuospatial extensions (the number of displayed elements) of the two subclasses (A and A') in order to activate the appropriate logical (or conceptual) comparison of the superordinate class (B) to its subordinate class (A) – the class-inclusion algorithm. In the negative priming adaptation of the class-inclusion task, adults and 10-year-old children performed test and control trials with three types of items: class-inclusion items, subclasses-comparison items, and control items (Borst et al., 2013b). Stimuli consisted of two rows of various geometric shapes of different colors separated by a horizontal line (e.g., eight green squares and four blue squares). Class-inclusion items (e.g., "More green squares than squares": yes or no?) required to compare the superordinate class (e.g., squares) to one of its subordinate classes (e.g., green squares). Subclasses-comparison items required to compare the number of elements in the two subclasses (e.g., "More green squares than blue squares"). On control items participants were required to judge whether all objects had the same given property (e.g., "Squares have the same color"). In the test trials, participants performed a typical class-inclusion item on the prime (in which inhibition of the comparison of the subordinate classes' extensions was needed) and then a subclasses-comparison item on the probe (in which the direct comparison of the two subclasses' extensions became the appropriate strategy, e.g., comparing the number of blue and green squares). In the control trials, participants performed a control item on the prime followed by a subclasses-comparison item on the probe. Critically, the strategy to be used on the prime was not related to the strategy to be used on the probe. Negative priming was reported for both children and adults: children and adults were slower to determine that there were more objects in one subordinate class than in the other after they successfully determine that there were more elements in the superordinate class than in one of the two subordinate classes. In addition, negative priming decreased with age. The results reported in this study extend the related findings of Perret et al. (2003) in school-aged children by showing (a) that young adults still need to inhibit the misleading perceptual strategy – i.e., the direct comparison of the subordinate classes – to reason about class inclusion and (b) that the efficiency of the inhibitory control needed in this specific task increases between fourth graders and young adults.

Another study from our lab has used a negative priming paradigm to demonstrate that inhibition is required in syllogistic reasoning (Moutier et al., 2006; Borst et al., 2013a). As in other negative priming studies, children performed test and control trials. Each trial consisted of two syllogisms with many words in common. In test trials, on the prime the validity of the syllogism was in contradiction with children's knowledge of the world (e.g., *All elephants are light*). Therefore, children had to inhibit their belief heuristic (e.g., elephants are heavy) to correctly judge the logical validity of the conclusion. On the probe, a syllogism was presented in which children's belief was congruent with the logical validity of the conclusion (e.g., *All elephants are light,* when the conclusion was not valid). Critically, the belief that was inhibited on the prime was congruent with the validity of the syllogism on the probe. On control trials, children solved neutral syllogisms in which the conclusion was neither unacceptable nor acceptable regarding the children's beliefs (e.g., *No students in the blue school are interested in sports*) followed, on the probe, by a syllogism in which the belief was congruent with the logical validity of the conclusion. As expected if inhibitory control is needed for syllogistic reasoning a negative priming effect was reported on the number of errors made by the participants: children committed more errors on the congruent syllogisms (probe items) when performed after syllogisms (prime items) in which beliefs and the validity of the conclusion interfered. Thus, as with the other logicomathematical tasks that we reviewed, syllogistic reasoning seems directly related to the ability to inhibit irrelevant strategies (or beliefs) in order to activate a logical algorithm.

Further studies are needed to investigate whether inhibitory control development during childhood and adolescence contributes to conceptual development in other cognitive domains than the ones we investigated (i.e., number, categorization, and reasoning).

FROM THE LAB TO THE CLASSROOM

Finally, beyond classical laboratory experimental situations, it seems that some systematic difficulties children have in resolving

problems in the classroom are also related to their difficulty to inhibit what they previously learned. For example, we investigated whether simple arithmetic word problems such as "Bill has 20 marbles. He has five more marbles than John. How many marbles does John have?" could remain challenging for children because they fail to inhibit the "add if more, subtract if less" misleading heuristic. Indeed, errors in this type of problems are characterized by adding the numbers instead of subtracting them or vice versa. Using a negative priming paradigm, we demonstrated that children and even adults must inhibit the "add if more, subtract if less" misleading strategy to solve simple arithmetic word problems in which the relational term ("more" or "less") is incongruent with the arithmetic operation to perform (Lubin et al., 2013). Thus, the increased efficiency to solve this type of problems from childhood to adulthood may be directly related to the gradual development of inhibitory control efficiency.

This new approach of cognitive development opens an avenue for designing pedagogical interventions (in line with Zelazo, 2006; Diamond et al., 2007; Houdé, 2007; Chevalier and Blaye, 2008; Diamond and Lee, 2011; Moriguchi, 2012) based on training the inhibition of heuristics (or reasoning biases). Previous studies have demonstrated that this type of pedagogical interventions not only improve logical reasoning to a greater extent than ones based solely on verbal logic *per se* (e.g., Houdé et al., 2000; Moutier and Houdé, 2003; Houdé, 2007; Cassotti and Moutier, 2010) but also help children in the classroom overcome systematic difficulties to a greater extent than traditional curricula (Lubin et al., 2012). Further studies much as the ones conducted on the effect of training working memory, another domain-general executive function (see e.g., Olesen et al., 2004), are needed to determine more precisely the effect of inhibitory control training on the development of the prefrontal cortex.

REFERENCES

Antell, S., and Keating, D. (1983). Perception of numerical invariance in neonates. *Child Dev.* 54, 695–701. doi: 10.2307/1130057

Apperly, I. A., and Carroll, D. J. (2009). How do symbols affect 3- to 4-year-olds' executive function? Evidence from a strategic reasoning task. *Dev. Sci.* 12, 1070–1082. doi: 10.1111/j.1467-7687.2009.00856.x

Baillargeon, R. (1995). "Physical reasoning in infancy," in *The Cognitive Neurosciences*, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 181–204.

Beck, S. R., Riggs, K. J., and Gorniak, S. L. (2009). Relating developments in children's counterfactual thinking and executive functions. *Think. Reason.* 15, 337–354. doi: 10.1080/13546780903135904

Bell, M., and Fox, N. (1992). The relations between frontal brain electrical activity and cognitive development during infancy. *Child Dev.* 63, 1142–1163. doi: 10.2307/1131523

Benson, J., Sabbagh, M. A., Carlson, S. M., and Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Dev. Psychol.* 49, 1615–1627. doi: 10.1037/a0031056

Berger, A., Tzur, G., and Posner, M. (2006). Infant brains detect arithmetic errors. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12649–12653. doi: 10.1073/pnas.0605350103

Bjorklund, D., and Harnishfeger, K. (1990). The resources construct in cognitive development: diverse sources of evidence and a theory of inefficient inhibition. *Dev. Rev.* 10, 48–71. doi: 10.1016/0273-2297(90)90004-N

Borst, G., Moutier, S., and Houdé, O. (2013a). "Negative priming in logicomathematical reasoning: the cost of blocking your intuition," in *New Approaches in Reasoning Research – Current Issues in Thinking and Reasoning*, eds W. De Neys and M. Osman (New York: Psychology Press), 34–50.

Borst, G., Pineau, A., Poirel, N., Cassotti, M., and Houdé, O. (2013b). Inhibitory control efficiency in a Piaget-like class-inclusion task in school-age children and adults: a developmental negative priming study. *Dev. Psychol.* 49, 1366–1374. doi: 10.1037/a0029622

Borst, G., Simon, G., Vidal, J., and Houdé, O. (2013c). Inhibitory control and visuospatial reversibility in Piaget's seminal number-conservation task: a high-density ERP study. *Front. Hum. Neurosci.* 7:920. doi: 10.3389/fnhum.2013.00920

Borst, G., Poirel, N., Pineau, A., Cassotti, M., and Houdé, O. (2012). Inhibitory control in number-conservation and class-inclusion tasks: a neo-Piagetian inter-tasks priming study. *Cogn. Dev.* 27, 283–298. doi: 10.1016/j.cogdev.2012.02.004

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Casey, B., Tottenham, N., Liston, C., and Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends Cogn. Sci.* 9, 104–110. doi: 10.1016/j.tics.2005.01.011

Cassotti, M., and Moutier, S. (2010). How to explain receptivity to conjunction-fallacy inhibition training. *Brain Cogn.* 72, 378–384. doi: 10.1016/j.bandc.2009.11.004

Chevalier, N., and Blaye, A. (2008). Cognitive flexibility in preschoolers: the role of representation activation and maintenance. *Dev. Sci.* 11, 339–353. doi: 10.1111/j.1467-7687.2008.00679.x

Daurignac, E., Houdé, O., and Jouvent, R. (2006). Negative priming in a numerical Piaget-like task in adults as evidenced by ERP. *J. Cogn. Neurosci.* 18, 730–736. doi: 10.1162/jocn.2006.18.5.730

De Neys, W., Lubin, A., and Houdé, O. (2014). The smart non-conserver: preschoolers detect their number conservation errors. *Child Dev. Res.* 2014:768186. doi: 10.1155/2014/768186

De Neys, W., Rossi, S., and Houdé, O. (2013). Bats, balls, and substitution sensitivity: cognitive misers are no happy fools. *Psychon. Bull. Rev.* 20, 269–273. doi: 10.3758/s13423-013-0384-5

Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics.* New York: Oxford University Press.

Demetriou, A. (ed.). (1988). *The Neo-Piagetian Theories of Cognitive Development.* Amsterdam: North-Holland.

Dempster, F., and Brainerd, C. (eds). (1995). *Interference and Inhibition in Cognition.* New York: Academic Press.

Diamond, A., Barnett, W., Thomas, J., and Munro, S. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Diamond, A., and Goldman-Rakic, P. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task. *Exp. Brain Res.* 74, 24–40. doi: 10.1007/BF00248277

Diamond, A. (1991). "Neuropsychological insights into the meaning of object concept development," in *The Epigenesis of Mind: Essays on Biology and Cognition*, eds S. Carey and R. Gelman (Hillsdale, NJ: Lawrence Erlbaum), 67–110.

Diamond, A. (1998). Understanding the A-not-B error: working memory vs. reinforced response, or active trace vs. latent trace. *Dev. Sci.* 1, 185–189. doi: 10.1111/1467-7687.00022

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529

Evans, J. (2003). In two minds: dual-process accounts of reasoning. *Trends Cogn. Sci.* 7, 454–459. doi: 10.1016/j.tics.2003.08.012

Gelman, R. (1972). Logical capacity of very young children. *Child Dev.* 43, 75–90. doi: 10.2307/1127873

Gelman, R. (1997). Constructing and using conceptual competence. *Cogn. Dev.* 12, 305–313. doi: 10.1016/S0885-2014(97)90002-2

Gelman, R., and Meck, E. (1983). Preschooler's counting. *Cognition* 13, 343–359. doi: 10.1016/0010-0277(83)90014-8

Gelman, R., Meck, E., and Merkin, S. (1986). Young children's numerical competence. *Cogn. Dev.* 1, 1–29. doi: 10.1016/S0885-2014(86)80021-1

Gopnik, A. (2012). Scientific thinking in young children: theoretical advances, empirical research, and policy education. *Science* 337, 1623–1627. doi: 10.1126/science.1223416

Gopnik, A., Meltzoff, A., and Kuhl, P. (1999). *The Scientist in the Crib*. New York: William Morrow and Cie.

Harnishfeger, K. (1995). "The development of cognitive inhibition: theories, definition, and research evidence," in *Interference and Inhibition in Cognition*, eds F. Dempster and C. Brainerd (New York: Academic Press), 176–204.

Houdé, O. (2000). Inhibition and cognitive development: object, number, categorization, and reasoning. *Cogn. Dev.* 15, 63–73. doi: 10.1016/S0885-2014(00)00015-0

Houdé, O. (2007). First insights on neuropedagogy of reasoning. *Think. Reason.* 13, 81–89. doi: 10.1080/13546780500450599

Houdé, O. (2014). "Cognitive development during infancy and early childhood across cultures," in *International Encyclopedia of the Social and Behavioral Sciences*, 2nd Edn, ed. J. D. Wright (Oxford: Elsevier Science).

Houdé, O., and Guichart, E. (2001). Negative priming effect after inhibition of number/length interference in a Piaget-like task. *Dev. Sci.* 4, 71–74. doi: 10.1111/1467-7687.00156

Houdé, O., Pineau, A., Leroux, G., Poirel, N., Perchey, G., Lanoë, C., et al. (2011). Functional MRI study of Piaget's conservation-of-number task in preschool and school-age children: a neo-Piagetian approach. *J. Exp. Child Psychol.* 110, 332–346. doi: 10.1016/j.jecp.2011.04.008

Houdé, O., and Tzourio-Mazoyer, N. (2003). Neural foundations of logical and mathematical cognition. *Nat. Rev. Neurosci.* 4, 507–514. doi: 10.1038/nrn1117

Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B., et al. (2000). Shifting from the perceptual brain to the logical brain: the neural impact of cognitive inhibition training. *J. Cogn. Neurosci.* 12, 721–728. doi: 10.1162/089892900562525

Inhelder, B., and Piaget, J. (1964). *The Early Growth of Logic in the Child*, A. Lunzer and D. Papert (trans.) (New York, NY: Routledge and Kegan Paul, original work published 1959).

Johnson, M. (2010). Interactive specialization: a domain-general framework for human functional brain development? *Dev. Cogn. Neurosci.* 1, 7–21. doi: 10.1016/j.dcn.2010.07.003

Kahneman, H. (2011). *Thinking Fast and Slow*. London: Allen Lane.

Lipton, J., and Spelke, E. (2003). Origins of number sense: large number discrimination in human infants. *Psychol. Sci.* 14, 396–401. doi: 10.1111/1467-9280.01453

Loosbroek, E., and Smitsman, A. (1990). Visual perception of numerosity in infancy. *Dev. Psychol.* 26, 916–922. doi: 10.1037/0012-1649.26.6.911.b

Lubin, A., Lanoë, C., Pineau, A., and Rossi, S. (2012). *Apprendre à inhiber: Une pédagogie innovante au service des apprentissages scolaires fondamentaux (mathématiques et orthographe) chez des élèves de 6 à 11 ans* [Inhibition training: an innovative pedagogy for academic learning (mathematics and spelling) in schoolchildren aged 6 to 11 years]. *Neuroéducation* 1, 55–84.

Lubin, A., Vidal, J., Lanoë, C., Houdé, O., and Borst, G. (2013). Inhibitory control is needed for the resolution of arithmetic word problems: a developmental negative priming study. *J. Educ. Psychol.* 105, 701–708. doi: 10.1037/a0032625

Mandler, J. (1988). How to build a baby: on the development of an accessible representational system. *Cogn. Dev.* 3, 113–136. doi: 10.1016/0885-2014(88)90015-9

Mehler, J., and Bever, T. (1967). Cognitive capacity of very young children. *Science* 158, 141–142. doi: 10.1126/science.158.3797.141

Moriguchi, Y. (2012). The effect of social observation on children's inhibitory control. *J. Exp. Child Psychol.* 113, 248–258. doi: 10.1016/j.jecp.2012.06.002

Moutier, S., and Houdé, O. (2003). Judgement under uncertainty and conjunction fallacy inhibition training. *Think. Reason.* 9, 185–201. doi: 10.1080/13546780343000213

Moutier, S., Plagne, S., Melot, A.-M., and Houdé, O. (2006). Syllogistic reasoning and belief-bias inhibition in school children. *Dev. Sci.* 9, 166–172. doi: 10.1111/j.1467-7687.2006.00476.x

Neill, W. T., Valdes, L. A., and Terry, K. M. (1995). "Selective attention and inhibitory control of cognition," in *Interference and Inhibition in Cognition*, eds F. N. Dempster and C. J. Brainerd (New York, NY: Academic Press), 207–261. doi: 10.1016/B978-012208930-5/50008-8

Olesen, P., Westerberg, H., and Klingberg, T. (2004). Increased prefrontal and parietal brain activity after training of working memory. *Nat. Neurosci.* 7, 75–79. doi: 10.1038/nn1165

O'Reilly, R. (1998). Six principles for biologically based computational models of cortical cognition. *Trends Cogn. Sci.* 11, 455–462. doi: 10.1016/S1364-6613(98)01241-8

Perret, P., Blaye, A., and Paour, J.-L. (2003). Respective contributions of inhibition and knowledge levels in class inclusion development: a negative priming study. *Dev. Sci.* 6, 283–286. doi: 10.1111/1467-7687.00284

Piaget, J. (1952). *The Child's Conception of Number*. London: Routledge and Kegan.

Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books. doi: 10.1037/11168-000

Piaget, J. (1983). "Piaget's theory," in *Handbook of Child Psychology*, Vol. 1, ed. P. H. Mussen (New York: Wiley), 103–128.

Poirel, N., Borst, G., Simon, G., Rossi, S., Cassotti, M., Pineau, A., et al. (2012). Number conservation is related to children's prefrontal inhibitory control: an fMRI study of a Piagetian task. *PLoS ONE* 7:e40802. doi: 10.1371/journal.pone.0040802

Siegler, R. (1999). Strategic development. *Trends Cogn. Sci.* 3, 430–435. doi: 10.1016/S1364-6613(99)01372-8

Siegler, R. S. (1996). *Emerging Minds: the Process of Change in Children's Thinking*. New York: Oxford University Press.

Simon, T. (1997). Reconceptualizing the origins of number knowledge: a non-numerical account. *Cogn. Dev.* 12, 349–372. doi: 10.1016/S0885-2014(97)90008-3

Simon, T. (1998). The foundations of numerical thinking in a brain without numbers. *Trends Cogn. Sci.* 3, 363–365. doi: 10.1016/S1364-6613(99)01383-2

Spelke, E. (2000). Core knowledge. *Am. Psychol.* 55, 1233–1243. doi: 10.1037/0003-066X.55.11.1233

Téglas, E., Vul, E., Girotto, V., Gonzales, M., Tenenbaum, J.-B., and Bonatti, L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science* 332, 1054–1059. doi: 10.1126/science.1196404

Tipper, S. P. (1985). The negative priming effect: inhibitory priming by ignored objects. *Q. J. Exp. Psychol. A* 37, 571–590. doi: 10.1080/14640748508400920

Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? *Q. J. Exp. Psychol. A* 54, 321–343. doi: 10.1080/713755969

Wright, I., Waterman, M., Prescott, H., and Murdoch-Eaton, D. (2003). A new Stroop-like measure of inhibitory function development: typical developmental trends. *J. Child Psychol. Psychiatry* 44, 561–575. doi: 10.1111/1469-7610.00145

Wynn, K. (1992). Addition and subtraction by human infants. *Nature* 358, 749–750. doi: 10.1038/358749a0

Wynn, K. (2000). Findings of addition and subtraction in infants are robust and consistent. *Child Dev.* 71, 1535–1536. doi: 10.1111/1467-8624.00245

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2014; accepted: 31 May 2014; published online: 18 June 2014.
Citation: Houdé O and Borst G (2014) Measuring inhibitory control in children and adults: brain imaging and mental chronometry. Front. Psychol. 5:616. doi: 10.3389/fpsyg.2014.00616
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Houdé and Borst. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 10 June 2014
doi: 10.3389/fpsyg.2014.00550

Executive functioning and reading achievement in school: a study of Brazilian children assessed by their teachers as "poor readers"

Pascale M. J. Engel de Abreu[1]*, Neander Abreu[2], Carolina C. Nikaedo[3], Marina L. Puglisi[4], Carlos J. Tourinho[1], Mônica C. Miranda[3], Debora M. Befi-Lopes[4], Orlando F. A. Bueno[3] and Romain Martin[1]

[1] *ECCS Research Unit, University of Luxembourg, Walferdange, Luxembourg*
[2] *Instituto de Psicologia, Universidade Federal da Bahia, Salvador, Brazil*
[3] *Departamento de Psicobiologia, Universidade Federal de São Paulo, São Paulo, Brazil*
[4] *Departamento de Fisioterapia Fonoaudiologia e Terapia Ocupacional, Universidade de São Paulo, São Paulo, Brazil*

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Caron Ann Campbell Clark, University of Oregon, USA
Helen St. Clair-Thompson, University of Newcastle, UK
Hannah Pimperton, University College London, UK

****Correspondence:***
Pascale M. J. Engel de Abreu, ECCS Research Unit, University of Luxembourg, B.P.2, L-7201 Walferdange, Luxembourg
e-mail: pascale.engel@uni.lu

This study examined executive functioning and reading achievement in 106 6- to 8-year-old Brazilian children from a range of social backgrounds of whom approximately half lived below the poverty line. A particular focus was to explore the executive function profile of children whose classroom reading performance was judged below standard by their teachers and who were matched to controls on chronological age, sex, school type (private or public), domicile (Salvador/BA or São Paulo/SP) and socioeconomic status. Children completed a battery of 12 executive function tasks that were conceptual tapping cognitive flexibility, working memory, inhibition and selective attention. Each executive function domain was assessed by several tasks. Principal component analysis extracted four factors that were labeled "Working Memory/Cognitive Flexibility," "Interference Suppression," "Selective Attention," and "Response Inhibition." Individual differences in executive functioning components made differential contributions to early reading achievement. The Working Memory/Cognitive Flexibility factor emerged as the best predictor of reading. Group comparisons on computed factor scores showed that struggling readers displayed limitations in Working Memory/Cognitive Flexibility, but not in other executive function components, compared to more skilled readers. These results validate the account that working memory capacity provides a crucial building block for the development of early literacy skills and extends it to a population of early readers of Portuguese from Brazil. The study suggests that deficits in working memory/cognitive flexibility might represent one contributing factor to reading difficulties in early readers. This might have important implications for how educators might intervene with children at risk of academic under achievement.

Keywords: executive function, reading, working memory, cognitive flexibility, selective attention, inhibition, poverty, learning difficulties

INTRODUCTION

Reading is a complex cognitive task that depends on a range of component skills. It is now well established that children's phonological awareness, letter-sound knowledge and broader oral language abilities play an important role in their reading development (Carroll et al., 2003; Muter et al., 2004; Nation et al., 2004; Rose, 2006; Nation et al., 2010; Fricke et al., 2013). More recently, executive functioning skills have been put forward as another crucial building block for literacy development. Children who struggle to read fluently or do not understand well what they read often have problems with their executive functions (Reiter et al., 2005; Sesma et al., 2009; Locascio et al., 2010; Pimperton and Nation, 2010, 2014). Much debate remains, however, regarding the exact nature and degree of executive functioning difficulties experienced by struggling readers.

The term "executive function" encompasses a collection of cognitive processes that people use to coordinate and control their thoughts and actions, particularly in novel situations, and that are crucial for higher-order problem solving and goal-directed behavior (Zelazo et al., 2008; Zelazo and Carlson, 2012). Executive functioning is often assessed by "complex tasks" such as the Tower of London or the Wisconsin Card Sorting Test that involve several lower-level executive functioning abilities. Studies using such complex executive function tasks generally report correlations with literacy (Hooper et al., 2002; Sesma et al., 2009). Recent theoretical models posit that in adults, different executive functions constitute distinct, yet related, components (Miyake et al., 2000; Friedman et al., 2008). There is also some evidence suggesting that executive functions represent a set of dissociable abilities in children, although the nature of these factors differs widely

across studies and developmental populations (Lehto et al., 2003; Senn et al., 2004; Huizinga et al., 2006; St. Clair-Thompson and Gathercole, 2006; Van der Sluis et al., 2007; Wiebe et al., 2008; Rose et al., 2011; Steele et al., 2012). Cognitive flexibility, working memory and inhibitory control are regarded by many as core components of executive functioning because they are relatively well defined conceptually, often emerge as dissociable constructs in factor-analytic models and have been shown to be implicated in performance on more complex executive function tasks (Baddeley, 1996; Roberts and Pennington, 1996; Rabbitt, 1997; Miyake et al., 2000; Lehto et al., 2003).

A concept closely related to executive function is attention and many descriptions of executive functioning also include subfunctions of attention (Klenberg et al., 2001; Manly et al., 2001; Breckenridge et al., 2013; Loher and Roebers, 2013). In an attempt to separate different executive function components through exploratory factor analyses in 7- to 12-year-old children, Klenberg et al. (2001) reported that inhibition, auditory attention, selective visual attention and fluency clustered into separate factors. Another exploratory factor analysis involving 11-year-olds identified a two factor structure: one associated with working memory and one with inhibition. The study also included measures of cognitive flexibility that failed, however, to relate to a third distinct executive factor (St. Clair-Thompson and Gathercole, 2006).

Working memory has been described as a cognitive system of multiple components that is used to store and work with information in mind for brief periods of time (Baddeley, 2000). Most theorists agree that working memory comprises mechanisms devoted to the maintenance of information over a short period of time, also referred to as short-term memory and processes responsible for executive control that regulate and coordinate those maintenance operations (Engle et al., 1999). Whereas so-called simple span tasks mainly tap into the short-term storage component of the working memory system, performance on complex span tasks, that involve the simultaneous processing and storage of information, has been argued to rely on both central executive resources and domain-specific short-term storage systems (Duff and Logie, 2001). Some studies show a large or even complete overlap between simple and complex span tasks of working memory (Alloway et al., 2006), and it has been claimed by some that both types of measures essentially tap into the same underlying cognitive process (Unsworth and Engle, 2007). There is some evidence for discrete verbal and visuo-spatial working memory components (Shah and Miyake, 1996; Friedman and Miyake, 2000; Jarvis and Gathercole, 2003; Kane et al., 2004). In children it has been shown that verbal and visuo-spatial working memory tasks can relate to the same underlying factor while at the same time accounting for unique variance in academic achievement (St. Clair-Thompson and Gathercole, 2006). Verbal and visuo-spatial working memory measures might thus reflect partly domain general mechanisms and partly the contribution of modality specific storage systems (Baddeley and Logie, 1999; St. Clair-Thompson and Gathercole, 2006).

Cognitive flexibility (also known as task switching or set shifting) refers to the ability to flexibly adapt thoughts or actions as demanded by the situation (Cragg and Nation, 2008). It is generally assessed by tasks that consist of different conditions and that require subjects to switch from one condition to another in response to an external cue. Inhibitory control denotes processes which are involved in suppressing dominant but irrelevant stimuli or responses (Nigg, 2000). Several subtypes of inhibition have been proposed (Barkley, 1997; Friedman and Miyake, 2004; Martin-Rhee and Bialystok, 2008; Nigg, 2000). For example, Martin-Rhee and Bialystok (2008) distinguished tasks of response inhibition that require to override habitual responses to univalent displays (e.g., Go/No-Go paradigm) from tasks of interference suppression that are based on bivalent displays in which two presented features indicate potentially conflicting responses (e.g., Flanker paradigm). Selective attention refers to the ability to focus on particular information and to screen out irrelevant stimuli. It is often assessed through visual search paradigms in which target objects or features must be located among other distracters (Manly et al., 2001; Scerif et al., 2004).

Emerging research supports the idea of the contribution made by executive functioning to reading development. Working memory has been the most frequently studied and numerous findings point toward a positive relationship between performance in working memory tasks and reading proficiency (Gathercole et al., 2006a,b; St. Clair-Thompson and Gathercole, 2006; Swanson and Sachse-Lee, 2001; Swanson et al., 2004, 2011; Welsh et al., 2010). Whereas verbal short-term memory tasks have been linked consistently to decoding skills, complex span tasks have been found to make significant contributions to reading comprehension (Swanson and Berninger, 1995; Engel de Abreu and Gathercole, 2012). Working memory has also been linked to other areas of academic learning. Children with low working memory capacity often make poor general academic progress, leading to the hypothesis that working memory might act as a bottleneck for learning (Gathercole and Alloway, 2008). Cognitive flexibility has also been associated with reading ability (Hooper et al., 2002; Van der Sluis et al., 2007; Welsh et al., 2010; Cartwright, 2012). In a study from the US, Welsh et al. (2010) found that preschoolers' cognitive flexibility skills predicted their decoding and word recognition abilities at the end of kindergarten. Similarly, Van der Sluis et al. (2007) have shown that cognitive flexibility was positively linked to word-reading efficiency in 9–12-year-old children from the Netherlands.

Few studies have investigated inhibition and selective attention in relation to reading. Inhibitory processes have been implicated in reading in some studies (De Beni et al., 1998; De Beni and Palladino, 2000) but not in others (Lan et al., 2011). Lan et al. (2011) explored inhibition, working memory and selective attention cross-culturally in preschool children from China and the US and found that selective attention was the most robust predictor for letter–word identification in both countries. In contrast, in a longitudinal study on 3–6-year-olds from the UK, Steele et al. (2012) did not find a relationship between the ability to select and sustain attention and word recognition a year later. There is some evidence that struggling readers display limitations in tasks of selective attention (Sireteanu et al., 2008; Romani et al., 2011). For example Casco et al. (1998) showed that 11–12-year-old children with the lowest performance on a selective attention task achieved significantly lower scores in reading fluency than children with the highest selective attention abilities.

Other research exploring the contribution of executive functioning to reading has focused on clinical groups. Reiter et al. (2005) found that compared to their typically developing peers, children with dyslexia manifested deficits on measures of verbal and visuo-spatial working memory, inhibition, planning, and cognitive flexibility. An increasing body of research also suggests that specific reading comprehension difficulties are linked to executive dysfunction (Nation et al., 1999; Cain, 2006; Sesma et al., 2009; Borella et al., 2010; Locascio et al., 2010; Pimperton and Nation, 2010). In a study from the US, Locascio et al. (2010) found that children with specific reading comprehension difficulties ("poor comprehenders") were impaired on tasks tapping planning and visuo-spatial working memory. Findings from the UK indicate, however, that poor comprehenders show domain specific working memory and inhibitory deficits that are restricted to the verbal domain (Pimperton and Nation, 2010).

In sum, differences in executive functioning have been reported in good and poor readers but it remains unclear which specific executive function components might be affected. Few studies have included a comprehensive battery of tasks tapping into various facets of executive functioning ability. Furthermore, previous studies have focused almost exclusively on English speaking children from the US or the UK. Little is known about the relationship between specific components of executive functioning and reading in other cultural and linguistic contexts.

THE PRESENT STUDY

This study was conducted in Brazil with typically developing children in the early primary school years. Children in Brazil learn to read and write in Portuguese, a Romance language that is spoken by approximately 200 million people world-wide. The Portuguese orthographic code is more transparent than the English one, although less transparent than other European languages such as German or Italian (Pinheiro, 1995). Despite major improvement over the last decade, many students in Brazil perform below expected levels of literacy. The latest figures from the OECD "Programme for International Student Assessment" (PISA) indicate that Brazil ranks 55 out of 65 countries on reading, with half of the country's students performing below the basic proficiency level (OECD, 2013). Constructivist teaching methods (also known as the "whole language" approach) represent the dominant approach to literacy instruction in Brazil (Abadzi, 2006; Belintane, 2006). This approach is based on the belief that children discover the alphabetic code spontaneously in the course of reading and writing, and stands in contrast to the skill-based phonics approach that is used widely in English-speaking countries (National Institute of Child Health and Human Development, 2000; Ehri et al., 2001).

Our study explored working memory, cognitive flexibility, inhibition and selective attention in a large sample of young children from a range of social backgrounds. Each of these executive function domains was assessed with multiple measures that were carefully selected from the cognitive neuroscience literature and that are widely used in research and clinical settings to measure processes related to executive functioning in children. The objective was to choose relatively simple tasks that conceptually tap into isolated executive function components. The first step toward understanding the nature of the contribution made by different components of executive functioning to reading is to explore whether these theoretically distinguishable executive functions are actually discernible as distinct factors in a population of Brazilian children from a range of demographic backgrounds. Notably, our sample was ethnically and socioeconomically diverse and approximately 50% of the children lived below the poverty line.

A major interest was to explore the executive function profile of children who were assessed by their teachers as low reading achievers but without a diagnosed learning disability. There is without a doubt much controversy over whether teachers can identify reliably those children with reading problems. In an educational system such as that in Brazil, teachers' judgment of children's level of achievement is, however, crucial because grade (i.e., school year) repetition is common practice and is primarily initiated by the school on the basis of teachers' judgment of children's levels of attainment (Bruns et al., 2011). Low achieving students are held in the same grade for an extra year rather than being promoted to a higher grade along with their age peers. In Brazil almost 25% of students in the first grade repeat a year (PREAL, 2009) with children from the poorest segments of society being most affected (Bruns et al., 2011). Costs associated with grade repetition in Brazil are among the highest in the world (OECD, 2011). According to a recent World Bank estimate, Brazil spends approximately 12% of its total basic education expenditure on grade repetition (Bruns et al., 2011). The problem of grade repetition is however not restricted to Brazil; approximately 32.2 million children in primary education worldwide repeat a grade (UNESCO, 2012). A major reason for grade repetition around the world is low levels of academic performance.

There is a general consensus that the ability to read is a fundamental educational skill that forms the basis for all future learning. Children need to be able to read well in order to engage fully in the classroom and learn. Today many students across the developing world have reading difficulties that can have tremendous long-term consequence for their academic achievement and later success in life. In Brazil, approximately one-third of third graders are not able to read more than isolated words and phrases or find specific information in text (PREAL, 2009). A better understanding of the cognitive profile of children with low reading achievement in the classroom is thus crucial for the early identification of children at risk of academic failure and to improve educational outcomes for disadvantaged children.

In summary, the purpose of this study is twofold. Firstly to explore the extent to which different executive function components relate to reading achievement as measured by teacher evaluation in the early school grades. Secondly, to shed light on the executive function profile that accompanies low reading achievement in general education classrooms in Brazil. Research considering various components of executive functioning in a single study in young children is limited. This is particularly true for children at increased risk of academic failure such as those from low-income homes who are often excluded from scientific studies. Our study addresses the following questions:

1. Do separable executive function components make differential contributions to reading achievement in early readers from Brazil?
2. Do Brazilian children whose classroom reading performance in decoding and reading comprehension is judged below standard by their teachers differ from children with average or good reading scores on measures of executive functioning?
3. What executive functioning components separate the performance of the low reading achievers from those with average or good reading scores and what is the predictive capacity of the identified components for classifying students as good or poor readers?

MATERIALS AND METHODS

SAMPLING PROCEDURE

Children were recruited from public (i.e., state) and private schools across two Brazilian states—Bahia (BA, Northeast) and São Paulo (SP, Southeast). A range of schools from neighborhoods of different socioeconomic status levels in the cities of Salvador (BA) and São Paulo (SP) were contacted. We avoided schools that were located in extremely poor or dangerous neighborhoods, charged very high school fees or were bilingual. In total, 17 primary schools (53 classrooms) agreed to participate, of which 11 were located in Salvador and 6 in São Paulo. The data was collected as part of a larger study on the effects of poverty on children's cognitive development. At the time the study was conducted, children in Brazil started their reading instruction in Year 1 of primary, when they were around 6 years of age.

Caregivers of children from 1° *Ano* (Year 1) and 2° *Ano* (Year 2) of the *Ensino Fundamental I* (primary education I) of the selected schools were invited to complete the *Questionário Brasileiro do Ambiente Infantil* (QBAI, Brazilian Questionnaire of Children's Background) that was designed for this study. It contains questions pertaining to early childhood experiences, information on the medical and developmental history of the child and demographic and socio-economic characteristics of the household. The nutritional status of each child was assessed using anthropometric measurements (height, weight, and mid-upper arm circumference) following the recommendations of the World Health Organization (2007). Children also completed a non-verbal reasoning/IQ test (Raven Progressive Colored Matrices, Raven et al., 1986). Exclusion criteria for participation in the study were: maternal alcohol or drug use reported during pregnancy; severe complications at pregnancy or birth; very premature births (less than 32 weeks of gestation) or very low birth weight (1500 g or less); neurological impairments, history of head injury or other significant medical problems; moderate or severe stunted growth (below -2 SD from median height-for-age of reference population); moderate or severe wasting (below -2 SD from median weight-for-height of reference population); developmental delays or intellectual disability; learning disorder; victims of abuse; scholarship holders (*bolsistas*); bilingualism and chronological age outside the expected range.

In total 482 caregivers were interviewed, of whom approximately half were sending their children to private schools. 82 children were not tested because they did not meet the inclusion criteria and 5 children were excluded due to missing data. Complete data was obtained on 395 children. Our aim was to recruit a sample of typically developing children. The developmental and medical history of a subsample of children had to be assessed further by a team of physicians and for some children missing background information had to be completed by additional interviews. This led to a further exclusion of 40 participants for the following reasons: significant medical concerns (e.g., low APGAR scores; eclampsia, $N = 13$); maternal alcohol or drug abuse during pregnancy ($N = 9$); very premature or very low birth weight ($N = 4$); undernutrition ($N = 2$); Raven's score below the 5th percentile ($N = 4$); learning disorder or significant sensory impairment ($N = 4$); victim of abuse or domestic violence ($N = 4$).

Three hundred and fifty-five participants fulfilled all inclusion criteria. Teachers of these children were asked to rate each child's word decoding and reading comprehension achievement on a scale from 0 (very bad) to 10 (very good). This format corresponds to the standard grading scale used in Brazilian schools. From this sample, children with scores at or below 5 in both word decoding and reading comprehension were selected. Our cutoff score for determining whether a child is a "poor reader" is based on common educational practice in Brazil where a score of 6 (sometimes 5) is generally considered the minimum passing grade. The number of children identified as poor readers (from a total $N = 355$) was 53 (13%). These poor readers were matched for chronological age, sex, school type (private or public), domicile (Salvador or São Paulo) and socioeconomic status with 53 children presenting satisfactory reading scores of 6 or above in both decoding and reading comprehension. For simplicity, the latter group is referred to as the group of "good readers."

PARTICIPANTS

Descriptive characteristics of the groups are reported in **Table 1**. All children lived in an urban setting, were monolingual in Portuguese, and had a mean age of 7 years and 6 months (ranging from 6 years and 4 months to 8 years and 10 months) with no significant difference in age [$t_{(104)} = 1.28$; $p = 0.20$] among the two groups. The information obtained from the QBAI allowed us to calculate for each child the score on the *Critério de Classificação Econômica Brasil* (CCEB, Brazilian Criteria for Economic Classification, ABEP, 2010). The CCEB is commonly used in Brazil to segment the population into different economic classes (eight classes ranging from A1=very high socioeconomic status to E=very low socioeconomic status). We also computed for each child the score on the International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom, 2010). The index is based on a meta-analysis by Ganzeboom et al. (1992) and was designed to capture the attributes of occupations that convert caregivers' education into income. The score can range from 16 (e.g., cleaner) to 90 (e.g., judge). The index was derived from caregiver responses on caregiver occupation and was based on the highest occupational level of either caregiver.

Key sample demographics were as follows: 57% were boys, 83% were frequenting public schools (free of charge), 53% lived in Salvador, and the majority (60%) of the sample fell in the social class C corresponding to gross mean household incomes

Table 1 | Descriptive characteristics of the sample according to group.

Measures	Poor readers ($N = 53$)				Good readers ($N = 53$)				Significance
	Freq.	Mean	*SD*	Range	Freq.	Mean	*SD*	Range	*p* level
Age (in months)	–	91.11	7.74	76–106	–	89.19	7.72	77–105	n.s.
Sex (% boys)	56.60	–	–	–	56.60	–	–	–	n.s.
School type (% public)	83.00	–	–	–	83.00	–	–	–	n.s.
City (% Salvador/BA)	52.80	–	–	–	52.80	–	–	–	n.s.
Economic class (CCEB)									
A1	3.80	–	–	–	3.80	–	–	–	n.s.
A2	7.50	–	–	–	7.50	–	–	–	n.s.
B1	5.70	–	–	–	5.70	–	–	–	n.s.
B2	13.20	–	–	–	13.20	–	–	–	n.s.
C1	26.40	–	–	–	26.40	–	–	–	n.s.
C2	34.00	–	–	–	34.00	–	–	–	n.s.
D	9.40	–	–	–	9.40	–	–	–	n.s.
International Socioeconomic Status Index		38.00	16.00	17–89		40.26	17.58	17–85	n.s.
Lenght of schooling (in months)		40.36	14.35	6–68		39.32	13.01	8–57	n.s.
Non-verbal reasoning (Raven, percentile)		46.49	23.13	11–99		60.24	25.06	11–99	<0.05
Academic achievement (out of 10)									
Decoding		3.85	1.31	1–5		8.21	1.42	6–10	<0.001
Reading compr.		3.83	1.20	1–5		8.19	1.49	6–10	<0.001
Writing		4.19	1.33	1–6.7		7.67	1.61	4–10	<0.001
Mathematics		4.62	1.66	1–10		8.12	1.57	4–10	<0.001
Oral language		5.24	1.41	1–9		8.34	1.53	5–10	<0.001
Science		5.36	1.77	1–10		8.26	1.36	5–10	<0.001
Composite		4.86	1.35	1–8		8.16	1.32	5–10	<0.001

CCEB: Critério de Classificação Econômica Brasil (Brazilian Criteria for Economic Classification, ABEP, 2010); Raven: Raven Colored Progressive Matrices. Reading compr: reading comprehension.

between R$ 933.00—1391.00 (∼US$ 393.00—585.00; ABEP, 2010). Groups were matched on these demographic variables; ratios across the two groups were therefore identical. No significant group differences emerged in terms of length of schooling [$t_{(104)} = 0.39$; $p = 0.70$] and the International Socioeconomic Status Index [$t_{(104)} = 0.69$; $p = 0.49$]. Approximately half of the children in each group (52% of the poor readers, 43% of the good readers) lived below the poverty line that was set at 50% of the median disposable income in Brazil (OECD, 2011). The sample was ethnically diverse: 50% of the children were multiracial, 25% were black and 25% were white. The good readers outperformed the poor readers on the measure of non-verbal reasoning [Raven: $t_{(104)} = 2.94$; $p < 0.05$].

As expected, significant group effects emerged for the classification measures decoding [$t_{(104)} = 16.45$, $p < 0.001$] and reading comprehension [$t_{(104)} = 16.53$, $p < 0.001$]. Significant group effects in favor of the good reading group also emerged on the non-classification measures of writing [$t_{(104)} = 12.14$, $p < 0.001$], mathematics [$t_{(104)} = 11.12$, $p < 0.001$], oral language [$t_{(104)} = 10.82$, $p < 0.001$], science [$t_{(104)} = 9.45$, $p < 0.001$], and on the scholastic achievement composite score [$t_{(104)} = 12.70$, $p < 0.001$]. Importantly, 85% of the poor readers achieved low writing scores (at or below 5); 72% had achieved failing mathematics scores, 59% were struggling with their oral language skills and 64% had difficulties in science. In the group of good readers percentages of children with scores at or below 5 were as follows: 9% for writing, 4% for mathematics, 2% for oral language and 2% for science.

TASK DEVELOPMENT

In Brazil, standardized tests that can be used to assess executive functioning in young children are scarce. The authors reviewed critically a large number of national and international instruments and discussed them with an expert panel composed of researchers, psychologists and teachers. Task selection was theory-driven. The material was carefully adapted or developed for the Brazilian context, and piloted on a Brazilian population. Task instructions from published English tests were translated into Brazilian Portuguese by a member of the research team (CJT) who is a native Brazilian and fluent in English. The translations together with the English originals were then revised by an expert panel of five independent assessors fluent in both Portuguese and English and the best features of the different revisions were retained. The measures were pretested and problematic items were further modified by the expert panel including the original translator. Reliability of instruments was established and is reported in the result section. A summary of the executive function tests that were used and the hypothesized executive function component that

Table 2 | Executive function measures selected for this study.

Hypothesized executive function component	Tasks
Cognitive flexibility	Duck task modified from the Dimensional Change Card Sort (Zelazo, 2006)
	Opposite worlds task from the Test of Everyday Attention for Children (Manly et al., 1998)
Working memory	Digit recall task of the Automated Working Memory Assessment (Alloway, 2007)
	Counting recall task of the Automated Working Memory Assessment (Alloway, 2007)
	Dot matrix task of the Automated Working Memory Assessment (Alloway, 2007)
	Odd-one-out task of the Automated Working Memory Assessment (Alloway, 2007)
Inhibition	O Mestre Mandou ("Simon says")
	Go/No-Go modified from Cragg and Nation (2008)
	Simon task
	Flanker task modified from the Attention Network Task from Rueda et al. (2004)
Selective attention	Map mission from the Test of Everyday Attention for Children (Manly et al., 1998)
	Sky search from the Test of Everyday Attention for Children (Manly et al., 1998)

they relate to are listed in **Table 2** and are described in detail below.

PROCEDURE

Informed written consent procedures were followed for all participants and the study was approved by the ethics committees of the University of Luxembourg and the Federal University of São Paulo, the *Hospital das Clínicas* of the School of Medicine of the University of São Paulo, the *Maternidade Climério de Oliveira* of the Federal University of Bahia, as well as the national Brazilian ethics committee *Comissão Nacional de Ética em Pesquisa* (CONEP, National Commission of Ethics in Research). Each child was assessed individually in a calm area of the school in two sessions that took place on different days and that lasted approximately 1 h each. Short breaks were used within sessions to maintain motivation. The measures were administered in a fixed sequence designed to vary the nature of the task demands across successive tests. Children received a sticker after completing different phases of the assessment and a diploma for their participation at the end of testing. They were tested by 8 research assistants who had all been trained by a member of the research team (PEdA). In total, children completed a battery of 19 tasks tapping executive functioning and other cognitive domains; the results on the 12 executive function tasks are reported here. Executive functioning was assessed with paper-and-pencil or computerized tasks.

For all measures, scores were converted to T-scores using the sample mean and standard deviation from the complete sample of 355 Brazilian children as a reference. The signs of the scores of variables on which low scores indicate better performance were inverted so that all positive scores represent superior performance.

MEASURES

Non-verbal reasoning

Children completed the *Raven Colored Progressive Matrices Test* (Raven et al., 1986) in which they have to complete a geometrical figure by choosing the missing piece among 6 possible drawings. Patterns increase progressively in difficulty and the test consisted of 36 items. Norms on a population of Brazilian children are available for this test (Angelini et al., 1999).

Cognitive flexibility

Two cognitive flexibility measures were administered: the *Duck Task* and the *Opposite Worlds* task. Both tasks contain different conditions and children have to switch from one condition to the other.

The *Duck Task* is a dimensional change card-sorting task that was modified from Zelazo (2006). Children have to sort bivalent test cards (red/blue; duck/flower) according to one specific rule (color or shape). The sorting rule changes across the task but the stimuli cards remain the same with each card representing the two dimensions at the same time. Two target cards (a blue duck and a red flower) are attached to sorting trays and remain visible throughout the task. Cards have to be placed facedown in one of the trays. Children are first told to sort the cards by shape ("shape game") and then by color ("color game"). In each case the experimenter explains the sorting rule and demonstrates two examples. The child then completes two practice trials with feedback followed by six experimental trials without feedback. In the next task, cards that contain an additional star sticker are introduced. Children are told that the star sticker cards need to be sorted by color whereas the cards without a star have to be sorted by shape ("shape-color game"). The experimenter demonstrates two examples (one with a star) and verifies verbally if the child understood the rules of the game. Children then complete two practice trials with feedback (one with a star). If the practice trials are failed the experimenter repeats the rules of the game and the child completes two further practice items with feedback. After these practice trials, the children are reminded of the rules of the game and then the experimental trials start. Children have to sort 24 cards with a rule reminder after 12 trials but no feedback. The majority of the cards (16) have to be sorted by shape; one-third of the trials (8) are star sticker trials. On the "shape game" and "color game" children scored at ceiling. The number of correctly sorted star sticker trials on the "shape-color game" was used as dependent variable in the analyses.

The *Opposite Worlds* task is part of the *Test of Everyday Attention for Children* (Manly et al., 1998). Children are presented

with stimulus sheets containing each a weaving path of the digits one and two. In the "same world" condition they have to follow the path and name the digits as quickly as possible in the conventional manner. In the "opposite world" condition they are asked to say "two" for the digit one and "one" for the digit two as they proceed along the path. The task starts with the "same world" condition, followed by two "opposite world" conditions and a final "same world" condition. The dependent variable used for these analyses was the sum of correct responses.

Working memory

Working memory was assessed with four sub-tests from the computer-based *Automated Working Memory Assessment* (AWMA, Alloway, 2007). The measures are verbal or visuo-spatial span tasks in which the number of items to be remembered increases progressively over successive blocks that contain six trials each. Testing stops after three errors in one block and the number of correctly recalled trials serves as the dependent variable.

Two verbal working memory measures—*Digit Recall* and *Counting Recall*—were administered. *Digit Recall* is a simple span task in which children have to repeat immediately sequences of spoken digits in the order that they were presented. In the *Counting Recall* task (a complex span task) children have to count and memorize the number of circles in pictures containing circles and triangles. At the end of each trial the number of circles in each picture has to be recalled in the right order.

Children also completed two measures of visuo-spatial working memory: The *Dot Matrix* and the *Odd-One-Out* tasks. In the simple span task *Dot Matrix*, children see a 4×4 matrix and a red dot that appears in different locations on the matrix. Children have to remember the sequence of the locations and recall them by tapping the squares of the empty matrix in the right order at the end of each trial. The *Odd-One-Out* task is a complex span task in which children are presented with arrays of three boxes with one shape in each. Two shapes are identical. Children have to identify the non-matching shape, remember its location in each array, and recall the localization of the odd shape when presented with an array of empty boxes at the end of the trial.

Inhibition

Response inhibition was assessed with two tasks ("*O Mestre Mandou*" and a *Go/No-Go* task) in which only certain conditions require a motor response while others must be ignored. Children also completed a *Simon* and a *Flanker* task of interference suppression. In these tasks, the features of bivalent displays either converge on a single response (creating congruent trials) or conflict by indicating different responses (creating incongruent trials).

In the "*O Mestre Mandou*" ("the master ordered") task, a Brazilian version of the children's game "Simon says," children stand opposite the experimenter who performs a series of physical actions accompanied by verbal commands (e.g., "touch your head"). Children have to imitate the actions of the experimenter if the command is prefaced with the phrase "*o mestre mandou*" but they must stand still for commands that do not begin with "*o mestre mandou.*" The experimenter performs all the actions irrespective of the instruction. In total, 16 trials are administered, of which 8 are non-imitation trials. The task is preceded by two practice trials with corrective feedback and children are reminded of the task rules after the first half of test trials. The dependent measure used for analyses is the sum of responses on the non-imitation trials that are coded as: 3 for no movement, 2 for wrong movement, 1 for partial imitation, and 0 for complete imitation.

The *Go/No-Go* task used was an adapted version of an English task by Cragg and Nation (2008). The task is presented on a laptop computer and consists of a background scene of a soccer goal and either a soccer ball (Go trials) or an American football ball (No-Go trials) that appears for 200 ms centrally near the bottom of the screen. Children are instructed to continuously press down the left mouse button (marked with a star) with the index finger of their dominant hand. When the soccer ball appears they are told that they have to shoot it by letting go of the star key and pressing the right mouse button as fast as possible with the same finger. When an American football ball appears they are told to keep their finger pressed down on the star key in order not to shoot the "funny looking" ball. Children first complete two blocks of 10 Go trials each. Next two mixed blocks (containing Go and No-Go trials) of 32 trials each are presented. No-Go stimuli occur on 25% of the trials. The dependent measure used for analyses was the percentage of correct responses in the mixed blocks. Go trials were scored as correct if the child released the star key and pressed the adjacent response key. No-go trials were scored as correct if the child continued pressing the star key.

The *Simon* and *Flanker* tasks were computer administered on a laptop. They were programmed and ran using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). Responses were recorded with two round colored response buttons (diameter of 2.5″), which were placed on the left and the right side next to the laptop keyboard. In the *Simon* task, green and yellow teddy bears ($2.75'' \times 2.56''$) appear on the left and the right side of the screen. Children have to press as quickly as they can the green response button if the teddy bear is green and the yellow button if the teddy bear is yellow. Half the trials are incongruent, so the colored teddy bear appears on the side opposite to the appropriate response button. The *Flanker* task was an adapted version of the Attention Network Task by Rueda et al. (2004). A horizontal row of five equally spaced yellow fish is presented ($3.35'' \times 0.39''$) and children have to indicate the direction of the central fish "Nemo" by pressing the corresponding left or right response buttons as fast as possible. On congruent trials (50% of all trials), the flanking fish are pointing in the same direction as the target, and on incongruent trials (50% of all trials), the distracters point in the opposite direction.

In both tasks, Simon and Flanker trials start with a fixation cross that appears in the middle of the screen for 1 s, followed by the stimulus for 5 s or until a response is made. Responses are followed by feedback and a 400-ms blank interval. Two blocks of 20 trials each have to be completed in which presentation of congruent and incongruent trials is randomized. Eight practice trials precede the experimental trials. If more than two errors occur on these trials, the instructions and the practice are repeated until the child reaches the criterion level. The dependent measures used for analyses were the reaction times (RTs) on incongruent trials

(excluding incorrect responses, RTs below 200 ms and RTs above 3 *SD* of individual means).

Selective attention

Two timed visual search tasks from the *Test of Everyday Attention for Children* (Manly et al., 1998) were administered: *Map Mission* and *Sky Search*.

In the *Map Mission* task children are presented with an A3 size city map with various distracters (e.g., small symbols of supermarket trolleys, cars…). They have to circle as many targets (small symbols of petrol stations) as possible within 1 min with a marker pen. In total 80 targets are presented. The dependent variable is the number of targets circled.

In the *Sky Search* task, children are given an A3-sheet with 128 paired spacecrafts of which 20 pairs are identical. They have to circle the identical pairs as quickly as possible with a marker pen. Next the children complete a motor control version of the task containing only the 20 target items. For both versions of the task, children have to mark a completion box when they are finished and timing is stopped. The motor control time-per-target score is subtracted from the initial time-per-target score leading to a Sky Search score that is relatively free from the impact of motor speed.

Classroom teacher ratings

Teachers were asked to rate each child's academic achievement during the school year on a scale from 0 to 10 in the following areas: *leitura* (reading): *decodificação* (decoding) and *compreensão* (comprehension); *escrita* (writing): *ortografia* (orthography), *redação* (text production), and *caligrafia* (handwriting); *matemática* (mathematics): *numeração* (numeracy), *contas* (basic arithmetic operations) and *compreensão de problemas* (problem solving); *linguagem oral* (oral language): *expressão* (expression) and *compreensão* (understanding); *ciências humanas e da natureza* (human and natural sciences): *ciências* (natural sciences), *história* (history), and *geografia* (geography). Composite scores were computed for writing, mathematics, oral language and human/natural sciences by averaging the different sub-scores in each domain. For each student the total level of achievement was also calculated.

RESULTS

DESCRIPTIVE STATISTICS AND CORRELATIONAL ANALYSES

The data did not present any missing values and none of the variables manifested severe departures from normality (Kline, 2005). Descriptive statistics for the non-verbal reasoning and executive function measures are provided in **Table 3**. Internal reliability estimates for the scores on the different measures were established for the complete sample ($N = 355$) using Cronbach's alpha. Reliability coefficients were in an acceptable range with reliability levels ranging from 0.60 to 0.93.

Zero-order correlation coefficients between age, non-verbal reasoning and the different executive function measures are reported in the upper triangle of **Table 4**. The lower triangle shows partial correlations controlling for chronological age (months). The overall pattern of relationship did not change when age was partialled out. As there exists a large overlap between fluid intelligence and executive functioning (Kyllonen and Christal, 1990; Engle et al., 1999; Conway et al., 2002; Colom et al., 2003; Kane et al., 2004; Engel de Abreu et al., 2010), non-verbal reasoning was not used as a covariate when exploring the relationship between the executive function components (Dennis et al., 2009).

Table 3 | Descriptive statistics for non-verbal reasoning and executive function scores ($N = 106$).

Measures	Mean	*SD*	Range	Cronbach's α
NON-VERBAL REASONING (PERCENTILE)				
Raven CPM	53.37	27.98	10–99	0.82
COGNITIVE FLEXIBILITY				
Duck task	46.03	9.52	34.19–63.42	0.87
Opposite worlds	44.93	9.79	20.06–62.74	0.66
WORKING MEMORY				
Digit recall	45.42	9.84	22.01–76.81	0.93
Counting recall	45.48	9.58	30.00–72.89	0.92
Dot matrix	46.58	9.55	23.08–68.95	0.91
Odd-one-out	46.47	9.88	30.88–75.74	0.91
INHIBITION				
Mestre mandou	49.42	11.14	20.86–65.77	0.60
Go/No-Go	50.11	10.38	26.92–69.34	0.73
Simon task	46.94	9.09	24.82–73.22	0.83
Flanker task	47.53	10.21	27.38–74.22	0.87
SELECTIVE ATTENTION				
Map mission	48.29	8.43	30.19–78.78	N/A
Sky search	47.55	9.69	30.99–75.95	N/A

Raven CPM: Raven Colored Progressive Matrices. With the exception of the Raven all scores are T scores. Cronbach's α was not computed on the timed selective attention measures.

As expected, non-verbal reasoning was significantly related to all the executive function measures, with the exception of the Go/No-Go task (r's ranging from 0.24 to 0.43). Within the areas of working memory and selective attention, measures correlated significantly with each other (r's ranging from 0.38 to 0.53) and correlations demonstrating convergent validity were higher than correlations demonstrating discriminant validity. A significant correlation was also obtained between the two cognitive flexibility measures ($r = 0.23$). These tasks also manifested moderate correlations with other executive function domains, especially with working memory (r's ranging between 0.26 and 0.37). For inhibition, a high correlation was obtained between the Simon and the Flanker tasks of interference suppression ($r = 0.64$) and a significant correlation emerged between the response inhibition measures Mestre mandou and Go/No-Go ($r = 0.25$). Notably, inter-correlations between measures of interference suppression and response inhibition were low (r's ranging from 0.07 to 0.20). Across executive function constructs, the working memory measures manifested the highest correlations, with the other executive function domains and the response inhibition measures the lowest.

COMPONENTS OF EXECUTIVE FUNCTIONING

The 12 executive function tasks were submitted to a principal component analysis with varimax rotation of the factor structure.

Table 4 | Correlations between age, non-verbal reasoning and executive functioning using Pearson's correlation coefficients ($N = 106$).

Measures	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Age	–	–	–	–	–	–	–	–	–	–	–	–	–	–
NON-VERBAL REASONING														
2. Raven CPM	−0.13	–	**0.35**	**0.34**	**0.43**	**0.44**	**0.47**	**0.45**	**0.26**	0.10	**0.40**	**0.36**	**0.35**	**0.26**
COGNITIVE FLEXIBILITY														
3. Duck task	−0.05	**0.35**	–	**0.23**	**0.26**	**0.35**	**0.26**	**0.30**	0.14	−0.07	0.13	0.04	0.02	0.09
4. Opposite worlds	0.04	**0.33**	**0.23**	–	**0.25**	**0.29**	**0.36**	**0.38**	**0.25**	**0.19**	**0.28**	0.17	**0.23**	**0.28**
WORKING MEMORY														
5. Digit recall	0.10	**0.41**	**0.26**	**0.26**	–	**0.52**	**0.39**	**0.38**	**0.31**	0.14	**0.23**	0.18	**0.25**	**0.22**
6. Counting recall	0.17	**0.41**	**0.34**	**0.29**	**0.53**	–	**0.38**	**0.49**	**0.25**	0.03	**0.28**	**0.25**	**0.23**	**0.33**
7. Dot matrix	**0.21**	**0.43**	**0.24**	**0.35**	**0.39**	**0.40**	–	**0.50**	**0.19**	0.18	**0.39**	**0.30**	**0.20**	0.10
8. Odd-one-out	**0.31**	**0.38**	**0.27**	**0.37**	**0.38**	**0.52**	**0.53**	–	**0.26**	0.03	**0.43**	**0.45**	0.11	0.20
INHIBITION														
9. Mestre mandou	0.10	**0.24**	0.14	**0.25**	**0.31**	**0.26**	**0.21**	**0.28**	–	**0.23**	0.17	**0.20**	**0.17**	0.15
10. Go/No-Go	0.17	0.10	−0.07	**0.20**	0.15	0.06	**0.21**	0.08	**0.25**	–	0.05	0.06	0.15	0.11
11. Simon task	0.17	**0.37**	0.12	**0.28**	**0.24**	**0.30**	**0.41**	**0.45**	0.18	0.07	–	**0.64**	0.17	0.22
12. Flanker task	0.10	**0.35**	0.03	0.17	0.18	**0.26**	**0.30**	**0.45**	**0.20**	0.07	**0.64**	–	**0.25**	0.19
SELECTIVE ATTENTION														
13. Map mission	**0.32**	**0.29**	0.00	**0.23**	**0.26**	**0.27**	**0.25**	**0.20**	0.19	**0.20**	**0.22**	**0.26**	–	**0.36**
14. Sky search	0.18	**0.24**	0.08	**0.28**	**0.23**	**0.35**	0.13	**0.24**	0.16	0.14	**0.24**	0.17	**0.39**	–

Raven CPM: Raven Colored Progressive Matrices. Upper triangle shows first-order correlations, and lower triangle shows correlations controlling for age in months. $p < 0.05$ are marked in boldface.

Four factors with eigenvalues above 1.00 were extracted, which accounted for 62.5% of the total variance. Factor loadings on the rotated matrix are listed in **Table 5**. A loading above 0.40 was used as a criterion for interpreting the factors. The working memory and cognitive flexibility measures loaded highly on Factor 1 (32.7%, factor loadings between 0.41 and 0.78). The interference suppression measures (Simon and Flanker) loaded on Factor 2 (10.8%, factor loadings of 0.85 and 0.87) with an additional moderate loading of the visuo-spatial working memory tasks (factor loadings of 0.46 and 0.57). Factor 3 (10.5%) included the subtests of selective attention (factor loadings of 0.75 and 0.83). The response inhibition measures loaded highly on Factor 4 (8.5%, factor loadings of 0.61 and 0.85). Notably, only the visuo-spatial working memory measures had loadings over 0.40 for more than one factor.

The extracted four components were labeled "Working Memory/Cognitive Flexibility" (Factor 1), "Interference Suppression" (Factor 2), "Selective Attention" (Factor 3), and "Response Inhibition" (Factor 4). For each participant factor scores produced by this solution were computed using the regression method and they were used as dependent measures for the subsequent analyses.

Table 5 | Factor loadings from principal component analysis.

Measures	Factor 1	Factor 2	Factor 3	Factor 4
	"Working Memory/Cognitive Flexibility"	***"Interference Suppression"***	***"Selective Attention"***	***"Response Inhibition"***
Duck task	**0.78**	−0.04	−0.11	−0.10
Opposite worlds	**0.41**	0.19	0.21	0.37
Digit recall	**0.62**	0.11	0.24	0.28
Counting recall	**0.70**	0.22	0.35	0.05
Dot matrix	**0.47**	**0.46**	0.00	0.35
Odd-one-out	**0.54**	**0.57**	0.10	0.15
Mestre mandou	0.27	0.11	0.06	**0.61**
Go/No-Go	−0.12	0.00	0.12	**0.85**
Simon task	0.13	**0.85**	0.14	0.05
Flanker task	0.00	**0.87**	0.14	0.05
Map mission	0.03	0.18	**0.75**	0.21
Sky search	0.18	0.07	**0.83**	0.02

Factor loadings above 0.40 are marked in boldface.

RELATIONSHIP BETWEEN EXECUTIVE FUNCTIONING COMPONENTS AND TEACHER RATINGS

Table 6 represents the partial correlation coefficients controlling for chronological age between the identified executive function factor structure and the different teacher ratings. The academic achievement ratings correlated strongly with each other (r's ranging from 0.80 to 0.98). Correlations between the decoding and reading comprehension ratings were high ($r = 0.98$).

Factor 1 correlated moderately to largely with all the teacher ratings of achievement (r's ranging from 0.29 to 0.43). Factor 2 was significantly related to reading, writing and mathematics (r's ranging from 0.20 to 0.29) and Factor 3 was linked significantly to ratings in reading (r's of 0.22 and 0.25) and oral language (r of 0.25). Weak associations emerged between Factor 4 and ratings of decoding and writing (r's of 0.22).

Considering the reading achievement ratings, the strongest correlations emerged with Factor 1 (r's of 0.35 and 0.36). These links were notably larger than the links for reading

Table 6 | Partial correlations (controlling for age in months) between the identified executive function factor structure and teacher ratings using Pearson's correlation coefficients ($N = 106$).

Teacher ratings	Factor 1	Factor 2	Factor 3	Factor 4	1	2	3	4	5	6	7
1. Decoding	**0.35***	**0.20**	**0.22**	**0.22**	–						
2. Reading compr.	**0.36***	**0.20**	**0.25***	0.18	**0.98**	–					
3. Writing	**0.29***	**0.25**	0.18	**0.22**	**0.86**	**0.90**	–				
4. Mathematics	**0.43***	**0.29***	0.18	0.14	**0.88**	**0.87**	**0.86**	–			
5. Oral language	**0.42***	0.18	**0.25**	0.13	**0.83**	**0.85**	**0.80**	**0.85**	–		
6. Sciences	**0.40***	0.19	0.19	0.17	**0.80**	**0.82**	**0.82**	**0.87**	**0.88**	–	
7. Composite	**0.41***	**0.23**	**0.22**	0.18	**0.91**	**0.92**	**0.91**	**0.94**	**0.93**	**0.96**	–

*Factor 1, "Working Memory/Cognitive Flexibility"; Factor 2, "Interference Suppression"; Factor 3, "Selective Attention"; Factor 4, "Response Inhibition"; Reading compr: reading comprehension. $p < 0.05$ are marked in boldface. *Correlation coefficients that remain significant after controlling for non-verbal reasoning.*

with the other executive function factors (*r*'s ranging from 0.18 to 0.25) and remained significant even after controlling for non-verbal reasoning (*r*'s ranging from 0.21 to 0.34).

PERFORMANCE OF THE GOOD AND POOR READERS ON EXECUTIVE FUNCTIONING COMPONENTS

A series of Analyses of Covariance were conducted with the executive function factor scores as dependent variables. After controlling for chronological age, significant group differences emerged on the Working Memory/Cognitive Flexibility factor [$F_{(1,\ 103)} = 9.29$; $p < 0.01$] with good readers outperforming poor readers (poor readers: $M = -0.28$, $SD = 1.01$; good readers: $M = 0.28$, $SD = 0.91$). This group effect remained significant even after controlling for non-verbal reasoning [$F_{(1,\ 102)} = 4.05$; $p < 0.05$]. The groups' performance was equivalently on the remaining factors.

A logistic regression analysis was conducted to predict reading group membership using Working Memory/Cognitive Flexibility as a predictor. A test of the full model against a constant-only model was statistically significant, indicating that the Working Memory/Cognitive Flexibility factor distinguished reliably between good and poor readers [$\chi^2_{(1)} = 8.64$; $p < 0.01$]. Prediction success overall was 61.3%, with 66% correctly classified for the group of poor readers and 57% for the group of good readers. The Wald criterion demonstrated that Working Memory/Cognitive Flexibility made a significant contribution to prediction ($p < 0.01$).

DISCUSSION

This research examined executive functioning and reading achievement in 6- to 8-year-old Brazilian children. Particular strengths of the study include the heterogeneity of the population sampled (drawn from a full range of socioeconomic backgrounds and reading achievement), the use of multiple measures tapping into different executive functioning components and the thorough group matching of participants on key socio-demographic factors. Findings showed, firstly, that in this population of children, individual differences in executive functioning components make differential contributions to early reading achievement. Secondly, that children whose classroom reading performance is judged below standard by their teachers demonstrate limitations in working memory/cognitive flexibility compared to more skilled readers.

The distinction between different executive function components fits well with findings from previous studies on adults (Robbins, 1996; Miyake et al., 2000; Friedman et al., 2008) and children (Lehto et al., 2003; Senn et al., 2004; Huizinga et al., 2006; St. Clair-Thompson and Gathercole, 2006; Van der Sluis et al., 2007) and is consistent with the multicomponential framework of executive functioning (Miyake et al., 2000). In this sample of young children from Brazil, the following four executive function components were identified: (1) Working Memory/Cognitive Flexibility, (2) Interference Suppression, (3) Selective Attention, and (4) Response Inhibition. Notably, measures of cognitive flexibility did not relate to a distinguishable underlying executive function construct but instead shared a common association with the working memory measures. This finding stands in contrast to studies on adults (Miyake et al., 2000) but is in line with other research on children, indicating that cognitive flexibility may be less differentiated from working memory in young children than in older children or adults (Senn et al., 2004; St. Clair-Thompson and Gathercole, 2006). It is worth noting that cognitive flexibility might well exist as a latent construct but might be difficult to identify in exploratory factor analyses because it might not account for a large amount of variance that is not shared with measures of working memory.

Another unexpected finding was that tasks of interference suppression and response inhibition were unrelated, indicating that these measures capture distinct aspects of inhibitory control. This extends previous evidence from Martin-Rhee and Bialystok (2008) and is consistent with the view that there are several distinguishable inhibitory components (Barkley, 1997; Nigg, 2000; Friedman and Miyake, 2004). Further, results showed that verbal and visuo-spatial working memory tasks as well as simple span (i.e., short-term memory) and complex span tasks of working memory related to the same underlying factor. This demonstrate that these measures rely, in part at least, on domain-general executive resources in young children. Visuo-spatial working memory tasks were additionally linked to measures of interference suppression. This finding fits well with the theoretical account on adults that the ability to deal with interference or conflict represents one key component of working memory capacity (Oberauer and Kliegl, 2001; Braver et al., 2007;

Hasher et al., 2007; Kane et al., 2007; Unsworth and Engle, 2007).

The results are consistent with previous research from English-speaking countries on independent contributions of discrete executive function components to children's academic achievement (St. Clair-Thompson and Gathercole, 2006), and extends those findings to a population of children from Brazil. The Working Memory/Cognitive Flexibility factor emerged as the best predictor of reading achievement and the magnitude of this relationship was considerably higher than the associations found between reading and other executive function components. It is notable that Working Memory/Cognitive Flexibility remained closely associated with the reading scores even when non-verbal reasoning was controlled. This result validates the account that working memory capacity provides a crucial building block for the development of early literacy skills (Swanson and Sachse-Lee, 2001; Gathercole et al., 2006a,b; St. Clair-Thompson and Gathercole, 2006; Welsh et al., 2010; Swanson et al., 2011) and shows that this relationship holds in early readers of Portuguese from Brazil.

Working memory/cognitive flexibility was also closely related to achievement in other academic domains, particularly mathematics. This finding is consistent with the view that working memory acts as a bottleneck for learning in that it supports general academic progress rather than the acquisition of skills and knowledge in specific domains (St. Clair-Thompson and Gathercole, 2006). According to Swanson and colleagues (Swanson and Saez, 2003; Swanson and Beebe-Frankenberger, 2004), working memory and scholastic achievement are related because greater working memory resources facilitate active maintainance of information and the integration of this with recent input and past knowledge. These represent key processes in academic learning. A related suggestion is that many classroom situations place heavy demands on the working memory system because children are required frequently to hold information in mind while engaging in effortful activities. Lengthy and complex classroom instructions or difficult task structures can lead to working memory overload in children with poor working memory function. This can result in task failure or abandonment, in other words, missed learning opportunities that negatively affect normal rates of learning (Gathercole et al., 2006b; Gathercole and Alloway, 2008).

Our study adds to existing evidence that struggling readers frequently display weaknesses in specific components of executive functioning. Compared to the good readers, children in the poor reading group had significantly lower scores on the Working Memory/Cognitive Flexibility factor. Unlike other authors, we did not find significant group differences on other executive function components (Reiter et al., 2005; Borella et al., 2010; Pimperton and Nation, 2010). The difference in findings could be due to the fact that previous studies focused almost exclusively on clinical populations of children with reading disorders such as dyslexia or specific reading comprehension difficulties. The present sample consisted of children without a diagnosed learning disability, drawn from typical classrooms but who had obtained low reading scores from their teachers.

It is worth noting that the focus of this study was on exploring the executive function profile of children whose classroom reading performance was judged below standard by their teachers and who were therefore at increased risk of grade repetition in Brazil. An obvious limitation of the study is that teacher ratings may be biased. It would be of interest if future studies would include standardized tests of reading achievement in a longitudinal research design. This would give a fuller appreciation of the nature of the relationship between executive functioning and reading.

This theoretical study has potential implications for practice and policy making. Learning to read is more than an educational skill. Low levels of literacy skills and living in poverty create a mutually reinforcing cycle that is difficult to break. The early identification of poor readers, together with remediation programmes that attempt to close gaps in achievement, are therefore crucial in order to counteract the impact of poverty on people's lives. Our study suggests that many students in Brazil might have fallen behind in their reading and struggle academically because of working memory limitations. Therefore teachers might want to assess whether underachieving students have working memory difficulties. Learning environments that prevent the overload of working memory resources might be a promising step toward counteracting early reading difficulties and subsequent school failure. Research from the UK has identified a number of methods of how to manage cognitive loads effectively in classroom settings (Gathercole et al., 2006b; Gathercole and Alloway, 2008). It remains to be seen whether such classroom-based approaches can enhance student learning in other cultural and educational settings. New research has also focused on supporting the development of working memory skills directly through targeted training programs (see Diamond and Lee, 2011, for a review). A range of activities have now been shown to improve children's working memory and might help children with poor academic progress overcome some of their learning difficulties (Holmes et al., 2009; Loosli et al., 2012; Alloway et al., 2013).

In conclusion, our findings indicate that distinct executive function components are predictive for individual differences in reading achievement in 6- to 8-year-old children. They also corroborate the notion that deficits in working memory/cognitive flexibility might represent one contributing factor to reading difficulties in early readers from Brazil.

ACKNOWLEDGMENTS

This research was supported by the following sources: National Research Fund (FNR) Luxembourg (Grant # CO09/LM/07, Pascale M. J. Engel de Abreu and Romain Martin); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil (Grant #400857/2010-3); Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brazil (Grant # 2010/11626-0 and Grant #2010/09185-5); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil; and Associação Fundo de Incentivo à Pesquisa (AFIP), Brazil. We would like to thank all the children, parents, teachers, and principals of the participating schools without whom this research would not have been possible. We are also grateful toward Ueslei Carneiro, Felipe Guedes, Lucas Carneiro, Manuela Sá and Adriana Rossi for assistance in data collection and to Anabela Cruz-Santos, Lucy Cragg, and

Carolina Toledo Piza for assistance and advice on test adaptation and design. Further, we wish to thank Dr. Med. Larissa de Freitas Rezende, Dr. Med. Maurício Costa de Abreu and Dr. Med. Maria Celeste Miranda Costa for their indispensable advice on medical exclusion criterion.

SUPPLEMENTARY MATERIAL

The Portuguese version for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00550/abstract

REFERENCES

Abadzi, H. (2006). *Efficient Learning for the Poor: Insights form the Frontier of Cognitive Neuroscience*. Washington, DC: The World Bank. doi: 10.1596/978-0-8213-6688-2

ABEP. (2010). *Critério de Classificação Econâmica Brasil*. Available online at: http://www.abep.org/new/criterioBrasil.aspx. (Accessed 22 January 2013).

Alloway, T. P. (2007). *Automated Working Memory Assessment*. London: Pearson Assessment.

Alloway, T. P., Bibile, V., and Lau, G. (2013). Computerized working memory training: can it lead to gains in cognitive skills in students? *Comput. Hum. Behav.* 29, 632–638. doi: 10.1016/j.chb.2012.10.023

Alloway, T. P., Gathercole, S. E., and Pickering, S. J. (2006). Verbal and visuo-spatial short-term and working memory in children: are they separable? *Child Dev.* 77, 1698–1716. doi: 10.1111/j.1467-8624.2006.00968.x

Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., and Duarte, J. L. M. (1999). *Matrizes Progressivas Coloridas de Raven: Escala Especial. Manual*. São Paulo: CETEPP.

Baddeley, A. D. (1996). Exploring the central executive. *Q. J. Exp. Psychol. A* 49, 5–28. doi: 10.1080/713755608

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/S1364-6613(00)01538-2

Baddeley, A. D., and Logie, R. (1999). "Working memory: the multiple component model," in *Models of Working Memory*, eds A. Miyake and P. Shah (New York, NY: Cambridge University Press), 28–61.

Barkley, R. A. (1997). *ADHD and the Nature of Self-Control*. New York, NY: Guilford Press.

Belintane, C. (2006). Reading and literacy in Brazil: a search beyond polarization. *Educação Pesquisa* 32, 261–277. doi: 10.1590/S1517-97022006000200004

Borella, E., Carretti, B., and Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *J. Learn. Disabil.* 43, 541–552. doi: 10.1177/0022219410371676

Braver, T. S., Gray, J. R., and Burgess, G. C. (2007). "Explaining the many varieties of working memory variation: dual mechanisms of cognitive control," in *Variation in Working Memory*, eds A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (New York, NY: Oxford University Press), 76–106.

Breckenridge, K., Braddick, O., and Atkinson, J. (2013). The organization of attention in typical development: a new preschool attention test battery. *Br. J. Dev. Psychol.* 31, 271–288. doi: 10.1111/bjdp.12004

Bruns, B., Evans, D., and Luque, J. (2011). *Achieving World-Class Education in Brazil: the Next Agenda*. Washington, DC: World Bank Publications. doi: 10.1596/978-0-8213-8854-9

Cain, K. (2006). Individual differences in children's memory and reading comprehension: an investigation of semantic and inhibitory deficits. *Memory* 14, 553–569. doi: 10.1080/09658210600624481

Carroll, J. M., Snowling, M. J., Hulme, C., and Stevenson, J. (2003). The development of phonological awareness in preschool children. *Dev. Psychol.* 39, 913–923. doi: 10.1037/0012-1649.39.5.913

Cartwright, K. B. (2012). Insights from cognitive neuroscience: the importance of executive function for early reading development and education. *Early Educ. Dev.* 23, 1–13. doi: 10.1080/10409289.2011.615025

Casco, C., Tressoldi, P. E., and Dellantonio, A. (1998). Visual selective attention and reading efficiency are related in children. *Cortex* 34, 531–546. doi: 10.1016/S0010-9452(08)70512-4

Colom, R., Flores-Mendoza, C., and Rebollo, I. (2003). Working memory and intelligence. *Pers. Individ. Dif.* 34, 33–39. doi: 10.1016/S0191-8869(02)00023-5

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., and Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence* 30, 163–183. doi: 10.1016/S0160-2896(01)00096-4

Cragg, L., and Nation, K. (2008). Go or no-go? Developmental improvements in the efficiency of response inhibition in mid-childhood. *Dev. Sci.* 6, 810–827. doi: 10.1111/j.1467-7687.2008.00730.x

De Beni, R., and Palladino, P. (2000). Intrusion errors in working memory tasks – are they related to reading comprehension ability? *Learn. Individ. Differ.* 12, 131–143. doi: 10.1016/S1041-6080(01)00033-4

De Beni, R., Palladino, P., Pazzaglia, F., and Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *Q. J. Exp. Psychol. A* 51, 305–320. doi: 10.1080/713755761

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., and Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *J. Int. Neuropsychol. Soc.* 15, 331–343. doi: 10.1017/S1355617709090481

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4–12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529

Duff, S. C., and Logie, R. H. (2001). Processing and storage in working memory span. *Q. J. Exp. Psychol. A* 54, 31–48. doi: 10.1080/02724980042000011

Engel de Abreu, P. M. J., Conway, A., and Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence* 38, 552–561. doi: 10.1016/j.intell.2010.07.003

Engel de Abreu, P. M. J., and Gathercole, S. E. (2012). Executive and phonological processes in second-language acquisition. *J. Educ. Psychol.* 104, 974–986. doi: 10.1037/a0028390

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128, 309–331. doi: 10.1037/0096-3445.128.3.309

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., and Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: evidence from the national reading panel's meta-analysis. *Read. Res. Q.* 36, 250–287. doi: 10.1598/RRQ.36.3.2

Fricke, S., Bowyer-Crane, C. A., Haley, A. J., Hulme, C., and Snowling, M. (2013). Efficacy of language intervention in the early years. *J. Child Psychol. Psychiatry* 54, 280–290. doi: 10.1111/jcpp.12010

Friedman, N. P., and Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *J. Exp. Psychol. Gen.* 129, 61–83. doi: 10.1037/0096-3445.129.1.61

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135. doi: 10.1037/0096-3445.133.1.101

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *J. Exp. Psychol. Gen.* 137, 201–225. doi: 10.1037/0096-3445.137.2.201

Ganzeboom, H. B. G. (2010). "A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with data from the ISSP 2002-2007," in *Paper presented at the Annual Conference of International Social Survey Programme* (Lisbon).

Ganzeboom, H. B. G., De Graaf, P., and Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Soc. Sci. Res.* 21, 1–56. doi: 10.1016/0049-089X(92)90017-B

Gathercole, S. E., and Alloway, T. P. (2008). *Working Memory and Learning: a Practical Guide*. London: Sage Press.

Gathercole, S. E., Alloway, T. P., Willis, C. S., and Adams, A. M. (2006a). Working memory in children with reading disabilities. *J. Exp. Child Psychol.* 93, 265–281. doi: 10.1016/j.jecp.2005.08.003

Gathercole, S. E., Lamont, E., and Alloway, T. P. (2006b). "Working memory in the classroom," in *Working Memory and Education*, ed S. Pickering (Elsevier Press), 219–240. doi: 10.1016/B978-012554465-8/50010-7

Hasher, L., Lustig, C., and Zacks, R. T. (2007). "Inhibitory mechanisms and the control of attention," in *Variation in Working Memory*, eds A. Conway, C. Jarrold, M. Kane, A. Miyake, and J. Towse (New York, NY: Oxford University Press), 227–249.

Holmes, J., Gathercole, S. E., and Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Dev. Sci.* 12, F9–F15. doi: 10.1111/j.1467-7687.2009.00848.x

Hooper, S., Swartz, C., Wakely, M., de Kruif, R., and Montgomery, J. (2002). Executive functions in elementary school children with and without problems in written expression. *J. Learn. Disabil.* 35, 37–68. doi: 10.1177/002221940203500105

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Agerelated change in executive function: developmental trends and a latent variables analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Jarvis, H. L., and Gathercole, S. E. (2003). Verbal and non- verbal working memory and achievements on national curriculum tests at 11 and 14 years of age. *Educ. Child Psychol.* 20, 123–140.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., and Engle, R. W. (2007). "Variation in working-memory capacity as variation in executive attention and control," in *Variation in Working Memory*, eds A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. Towse (New York, NY: Oxford University Press), 21–48.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., and Engle, R. W. (2004). The generality of working memory capacity: a latent variable approach to verbal and visuo-spatial memory span and reasoning. *J. Exp. Psychol. Gen.* 133, 189–217. doi: 10.1037/0096-3445.133.2.189

Klenberg, L., Korkman, M., and Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year old Finnish children. *Dev. Neuropsychol.* 20, 407–428. doi: 10.1207/S15326942DN2001_6

Kline, R. B. (2005). *Principles and Practice of Structural Equational Modeling*, *2nd Edn*. New York, NY: Guilford Press.

Kyllonen, P. C., and Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence* 14, 389–433. doi: 10.1016/S0160-2896(05)80012-1

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., and Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: a crosscultural analysis of chinese and american preschoolers. *J. Exp. Child Psychol.* 108, 677–692. doi: 10.1016/j.jecp.2010.11.001

Lehto, J. E., Juujärvi, Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Locascio, G., Mahone, E. M., Eason, S. H., and Cutting, L. E. (2010). Executive dysfunction among children with reading comprehension deficits. *J. Learn. Disabil.* 43, 441–454. doi: 10.1177/0022219409355476

Loher, S., and Roebers, C. M. (2013). Executive functions and their differential contribution to sustained attention in 5- to 8-year-old children. *J. Educ. Dev. Psychol.* 3, 51–63. doi: 10.5539/jedp.v3n1p51

Loosli, S. V., Buschkuehl, M., Perrig, W. J., and Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychol.* 18, 62–78. doi: 10.1080/09297049.2011.575772

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P., and Robertson, I. H. (2001). The differential assessment of children's attention: the Test of Everyday Attention for Children (TEA-Ch), normative sample data and ADHD performance. *J. Child Psychol. Psychiatry* 42, 1065–1081. doi: 10.1111/1469-7610.00806

Manly, T., Robertson, I. H., Anderson, V., and Nimmo-Smith, I. (1998). *Test of Everyday Attention for Children*. London: Pearson Assessment.

Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism* 11, 81–93. doi: 10.1017/S1366728907003227

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., and Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Muter, V., Hulme, C., Snowling, M. J., and Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Dev. Psychol.* 40, 665–681. doi: 10.1037/0012-1649.40.5.665

Nation, K., Adams, J. W., Bowyer-Crane, C. A., and Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *J. Exp. Child Psychol.* 73, 139–158. doi: 10.1006/jecp.1999.2498

Nation, K., Clarke, P., Marshall, C. M., and Durand, M. (2004). Hidden language impairments in children: parallels between poor reading comprehension and specific language impairment? *J. Speech Lang. Hear. Res.* 47, 199–211. doi: 10.1044/1092-4388(2004/017)

Nation, K., Cocksey, J., Taylor, J. S. H., and Bishop, D. V. M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *J. Child Psychol. Psychiatry* 51, 1031–1039. doi: 10.1111/j.1469-7610.2010.02254.x

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching Children to Read: an Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction (NIH Publication No. 00-4769)*. Washington, DC: U.S. Government Printing Office.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychol. Bull.* 126, 220–246. doi: 10.1037/0033-2909.126.2.220

Oberauer, K., and Kliegl, R. (2001). Beyond resources - formal models of complexity effects in age differences in working memory. *Eur. J. Cogn. Psychol.* 13, 187–215. doi: 10.1080/09541440042000278

OECD. (2011). *Society at a Glance 2011 - OECD Social Indicators, OECD publishing*, Available online at: www.oecd.org/els/social/indicators/SAG

OECD. (2013). *PISA 2012 Results in Focus: What 15-Year-Olds Know and What They can do With What They Know*. OECD Publishing. Available online at: http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

Pimperton, H., and Nation, K. (2010). Suppressing irrelevant information from working memory: evidence for domain specific deficits in poor comprehenders. *J. Mem. Lang.* 62, 380–391. doi: 10.1016/j.jml.2010.02.005

Pimperton, H., and Nation, K. (2014). Poor comprehenders in the classroom: teacher ratings of behavior in children with poor reading comprehension and its relationship with individual differences in working memory. *J. Learn. Disabil.* 47, 199–207. doi: 10.1177/0022219412454172

Pinheiro, A. M. V. (1995). Reading and spelling development in Brazilian Portuguese. *Read. Writ.* 7, 111–138. doi: 10.1007/BF01026950

PREAL. (2009). *Overcoming Inertia: a Report Card on Education in Brazil*. PREAL and the Lemann Foundation.

Rabbitt, P. (1997). "Methodologies and models in the study of executive function," in *Methodology of Frontal and Executive Function*, ed P. Rabbitt (East Sussex: Psychology Press Publishers), 1–38.

Raven, J. C., Court, J. H., and Raven, J. (1986). *Colored Progressive Matrices*. London: H. K. Lewis.

Reiter, A., Tucha, O., and Lange, K. W. (2005). Executive functions in children with dyslexia. *Dyslexia* 11, 116–131. doi: 10.1002/dys.289

Robbins, T. W. (1996). Dissociating executive functions of the prefrontal cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1463–1470. doi: 10.1098/rstb.1996.0131

Roberts, R. J., and Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Dev. Neuropsychol.* 12, 105–126. doi: 10.1080/87565649609540642

Romani, C., Tsouknida, E., and di Betta, A. M., Olson A. (2011). Reduced attentional capacity, but normal processing speed and shifting of attention in developmental dyslexia: evidence from a serial task. *Cortex* 47, 715–733. doi: 10.1016/j.cortex.2010.05.008

Rose, J. (2006). *Independent Review of the Teaching of Early Reading: Final Report*. Available online at: http://www.education.gov.uk/publications/eOrderingDownload/0201-2006pdf-EN-01.pdf. (Accessed 02 December 2013).

Rose, S. A., Feldman, J. F., and Jankowski, J. J. (2011). Modeling a cascade of effects: the role of speed and executive functioning in preterm/full-term differences in academic achievement. *Dev. Sci.* 14, 1161–1175. doi: 10.1111/j.1467-7687.2011.01068.x

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Scerif, G., Cornish, K., Wilding, J., Driver, J., and Karmiloff-Smith, A. (2004). Visual search in typically developing toddlers and toddlers with fragile X or Williams syndrome. *Dev. Sci.* 7, 116–130. doi: 10.1111/j.1467-7687.2004.00327.x

Senn, T. E., Espy, K. A., and Kaufmann, P. M. (2004). Using path analysis to understand executive function organization in preschool children. *Dev. Neuropsychol.* 26, 445–464. doi: 10.1207/s15326942dn2601_5

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., and Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychol.* 15, 232–246. doi: 10.1080/09297040802220029

Shah, P., and Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: an individual differences approach. *J. Exp. Psychol. Gen.* 125, 4–27. doi: 10.1037/0096-3445.125.1.4

Sireteanu, R., Goebel, C., Goertz, R., Werner, I., Nalewajko, M., and Thiel, A. (2008). "Impaired serial visual search in children with developmental dyslexia," in *Learning, Skill Acquisition, Reading, and Dyslexia*, Vol. 1145, eds G. F. Eden and D. L. Flower (Oxford: Blackwell), 199–211.

St. Clair-Thompson, H. L., and Gathercole, S. E. (2006). Executive functions and achievements on national curriculum tests: shifting, updating, inhibition, and working memory. *Q. J. Exp. Psychol.* 59, 745–759. doi: 10.1080/17470210500162854

Steele, A., Karmiloff-Smith, A., Cornish, K. M., and Scerif, G. (2012). The multiple sub-functions of attention: differential developmental gateways to literacy and numeracy. *Child Dev.* 83, 2028–2041. doi: 10.1111/j.1467-8624.2012.01809.x

Swanson, H. L., and Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for math disabilities. *J. Educ. Psychol.* 96, 471–491. doi: 10.1037/0022-0663.96.3.471

Swanson, H. L., and Berninger, V. (1995). The role of working memory in skilled and less skilled readers' comprehension. *Intelligence* 21, 83–108. doi: 10.1016/0160-2896(95)90040-3

Swanson, H. L., Orosco, M. J., Lussier, C. M., Gerber, M. M., and Guzman-Orth, D. A. (2011). The influence of working memory and phonological processing on English language learner children's bilingual reading and language acquisition. *J. Educ. Psychol.* 103, 838–856. doi: 10.1037/a0024578

Swanson, H. L., and Sachse-Lee, C. (2001). A subgroup analysis of working memory in children with reading disabilities: domain-general or domain-specific deficiency? *J. Learn. Disabil.* 34, 249–263. doi: 10.1177/002221940103400305

Swanson, H. L., and Saez, L. (2003). "Memory difficulties in children and adults with learning disabilities," in *Handbook of Learning Disabilities*, eds H. L. Swanson, S. Graham, and K. R. Harris (New York: Guilford), 182–198.

Swanson, H. L., Saez, L., Gerber, M., and Leafstedt, J. (2004). Literacy and cognitive functioning in bilingual and nonbilingual children at or not at risk for reading disabilities. *J. Educ. Psychol.* 96, 3–18. doi: 10.1037/0022-0663.96.1.3

UNESCO. (2012). *Global Education Digest 2012: Opportunities Lost: the Impact of Grade Repetition and Early School Leaving*. Available online at: http://www.uis.unesco.org/Education/Documents/ged-2012-en.pdf. (Accessed 12 November 2013).

Unsworth, N., and Engle, R. W. (2007). On the division of short-term and working memory: an examination of simple and complex spans and their relation to higher-order abilities. *Psychol. Bull.* 133, 1038–1066. doi: 10.1037/0033-2909.133.6.1038

Van der Sluis, S., de Jong, P. F., and van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence* 35, 427–449. doi: 10.1016/j.intell.2006.09.001

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., and Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low income families. *J. Educ. Psychol.* 102, 43–53. doi: 10.1037/a0016738

Wiebe, S. A., Espy, K. A., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575

World Health Organization. (2007). *WHO Child Growth Standards: Methods and Development: Head Circumference-for-Age, Arm Circumference-for-Age, Triceps Skinfold-for-Age and Subscapular Skinfold-for-Age*. Hong Kong: WHO.

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Zelazo, P. D., and Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: development and plasticity. *Child Dev. Perspect.* 6, 354–360. doi: 10.1111/j.1750-8606.2012.00246.x

Zelazo, P. D., Carlson, S. M., and Kesek, A. (2008). "Development of executive function in childhood," in *Handbook of Developmental Cognitive Neuroscience, 2nd Edn.,* eds C. A. Nelson and M. Luciana (Cambridge, MA: MIT Press), 553–574.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2014; accepted: 18 May 2014; published online: 10 June 2014.
*Citation: Engel de Abreu PMJ, Abreu N, Nikaedo CC, Puglisi ML, Tourinho CJ, Miranda MC, Befi-Lopes DM, Bueno OFA and Martin R (2014) Executive functioning and reading achievement in school: a study of Brazilian children assessed by their teachers as "poor readers." Front. Psychol. **5**:550. doi: 10.3389/fpsyg.2014.00550*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Engel de Abreu, Abreu, Nikaedo, Puglisi, Tourinho, Miranda, Befi-Lopes, Bueno and Martin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 17 June 2014
doi: 10.3389/fpsyg.2014.00599

Predictors of early growth in academic achievement: the head-toes-knees-shoulders task

Megan M. McClelland[1]*, Claire E. Cameron[2], Robert Duncan[1], Ryan P. Bowles[3], Alan C. Acock[1], Alicia Miao[1] and Megan E. Pratt[1]

[1] Human Development and Family Sciences, Oregon State University, Corvallis, OR, USA
[2] Center for Advanced Study of Teaching and Learning, University of Virginia, Charlottesville, VA, USA
[3] Department of Human Development and Family Studies, Michigan State University, East Lansing, MI, USA

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Yukio Maehara, Kyoto University, Japan
Sarah Loher, University of Bern, Switzerland

***Correspondence:**
Megan M. McClelland, Human Development and Family Sciences, Oregon State University, 245 Hallie E Ford Center for Healthy Children and Families, Corvallis, OR 97331-8687, USA
e-mail: megan.mcclelland@oregonstate.edu

Children's behavioral self-regulation and executive function (EF; including attentional or cognitive flexibility, working memory, and inhibitory control) are strong predictors of academic achievement. The present study examined the psychometric properties of a measure of behavioral self-regulation called the Head-Toes-Knees-Shoulders (HTKS) by assessing construct validity, including relations to EF measures, and predictive validity to academic achievement growth between prekindergarten and kindergarten. In the fall and spring of prekindergarten and kindergarten, 208 children (51% enrolled in Head Start) were assessed on the HTKS, measures of cognitive flexibility, working memory (WM), and inhibitory control, and measures of emergent literacy, mathematics, and vocabulary. For construct validity, the HTKS was significantly related to cognitive flexibility, working memory, and inhibitory control in prekindergarten and kindergarten. For predictive validity in prekindergarten, a random effects model indicated that the HTKS significantly predicted growth in mathematics, whereas a cognitive flexibility task significantly predicted growth in mathematics and vocabulary. In kindergarten, the HTKS was the only measure to significantly predict growth in all academic outcomes. An alternative conservative analytical approach, a fixed effects analysis (FEA) model, also indicated that growth in both the HTKS and measures of EF significantly predicted growth in mathematics over four time points between prekindergarten and kindergarten. Results demonstrate that the HTKS involves cognitive flexibility, working memory, and inhibitory control, and is substantively implicated in early achievement, with the strongest relations found for growth in achievement during kindergarten and associations with emergent mathematics.

Keywords: executive function, self-regulation, academic achievement, early childhood, measurement

INTRODUCTION

Self-regulation has been established as a key mechanism associated with a variety of outcomes including school readiness (Blair and Razza, 2007; McClelland et al., 2007a; Morrison et al., 2010), academic achievement during childhood and adolescence (McClelland et al., 2006; Cameron Ponitz et al., 2009; Duckworth et al., 2010; Li-Grining et al., 2010), and long-term health and educational outcomes (Moffitt et al., 2011; McClelland et al., 2013). Experts from diverse disciplines agree that self-regulation has important implications for individual health and well-being starting early in life (Geldhof et al., 2010; McClelland et al., 2010). Moreover, the behavioral aspects of self-regulation may be especially important for academic and school success (McClelland et al., 2007a; Cameron Ponitz et al., 2009; McClelland and Cameron, 2012). Given the multiple cognitive components involved in behavioral self-regulation, such as cognitive flexibility, working memory, and inhibitory control, measuring these skills during early childhood is challenging (Carlson, 2005; Cameron Ponitz et al., 2008; Caughy et al., 2014), and until recently, there have been few reliable and valid measures of these skills. Even fewer studies are able to address how well individual measures predict achievement growth over this significant developmental period or whether growth in behavioral measures are associated with growth in learning during the transition to kindergarten. The present study examined how a structured observation of behavioral self-regulation, the Head-Toes-Knees-Shoulders task (HTKS), was related to traditional executive function (EF) measures of cognitive flexibility, working memory, and inhibitory control. We also tested the predictive validity of these direct assessments for growth in academic achievement over four time points between preschool and kindergarten.

DEFINITIONS OF BEHAVIORAL SELF-REGULATION AND EXECUTIVE FUNCTION

Children's self-regulation of their cognitions, emotions, and behavior is critical for their success throughout the school trajectory and in adulthood (Zelazo and Müller, 2002; Baumeister and Vohs, 2004; Blair and Razza, 2007; McClelland et al., 2007a, 2013; Rimm-Kaufman et al., 2009). Different disciplines have examined self-regulation and related constructs using a variety of terms.

For example, scholars in the field of personality have used self-control to describe a set of skills similar to self-regulation and often refer to the integration of various self-control processes (Zimmerman, 2000; Eisenberg et al., 2014). And in the study of temperament, the construct of effortful control includes aspects of attentional focusing, inhibitory control, and regulating emotions, which are similar to self-regulation although temperament does not incorporate working memory (McClelland et al., 2010). In developmental psychology, self-regulation is a broad term that includes both top-down planning processes (e.g., executive functions or EF) and bottom-up regulation of more reactive impulses (Zelazo and Cunningham, 2007; Blair and Raver, 2012).

EF is a well-known construct originating in cognitive psychology that includes attentional or cognitive flexibility, working memory, and inhibitory control, which enables individuals to plan, organize, and problem-solve as well as to manage their impulses (Best and Miller, 2010). We have defined behavioral self-regulation as deliberately applying multiple component processes of attentional or cognitive flexibility, working memory, and inhibitory control to overt, socially contextualized behaviors like remembering to raise one's hand and waiting to be called upon instead of shouting out an answer (McClelland et al., 2007b; Cameron Ponitz et al., 2008; Morrison et al., 2010). Thus, whereas EF processes have typically been examined in terms of cognitive development, using materials and responses appropriate to the laboratory, behavioral self-regulation can be defined as the outward manifestation of those EF processes in adaptive, real-world behaviors (Cameron Ponitz et al., 2009; McClelland and Cameron, 2012). Throughout this paper we broadly refer to the set of contextualized, ecologically-relevant cognitive and behavioral processes as behavioral and use EF to refer specifically to individual cognitive components of attentional or cognitive flexibility, working memory, and inhibitory control. Whether a behavioral self-regulation measure is distinct from traditional EF measures in predicting academic achievement is one aim of this study.

The integration of EF into ecologically-relevant behaviors is critical for meeting school- and task-related demands and for successfully navigating early learning environments (McClelland and Cameron, 2012). For example, research indicates that behavioral self-regulation robustly contributes to achievement after controlling for initial achievement levels and other socio-demographic variables such as child IQ, age, ethnicity, and parent education level (Duncan et al., 2007; von Suchodoletz et al., 2009). In one recent study, a child with one standard deviation higher parent ratings of attention and persistence at age 4 had 49% higher odds of completing college by age 25 (McClelland et al., 2013). In another investigation, children with strong behavioral self-regulation in preschool had greater school age achievement after controlling for child IQ (von Suchodoletz et al., 2009).

The distinct roles played by the three individual EF components (attentional or cognitive flexibility, working memory, and inhibitory control) in regulating behavior is still debated (Barkley, 1997; Bronson, 2000; Müller et al., 2006). Attentional or cognitive flexibility allows children to shift focus and pay attention to new details, while simultaneously ignoring environmental distractions (Barkley, 1997; Rothbart and Posner, 2005). It may form the foundation for behavioral self-regulation and problem-solving (Zelazo and Müller, 2002; Rothbart and Posner, 2005; Rueda et al., 2005). Working memory allows children to remember and follow directions and helps them plan solutions to a problem (Gathercole and Pickering, 2000; Kail, 2003), and inhibitory control helps children stop one response in favor of a more adaptive behavior (Dowsett and Livesey, 2000; Carlson and Moses, 2001; Rennie et al., 2004).

Many measures of EF for young children produce a binary (pass/fail) distribution, which is consistent with Diamond et al. (2002) conceptualization of when children can keep track of multiple rules. In young children this depends on their ability to inhibit their initial impulse long enough to remember the rule and then give the correct response. Keeping track of and manipulating multiple rules (utilizing working memory) while also inhibiting initial impulses and activating an unnatural response is especially challenging for children. Our conceptualization of behavioral self-regulation is based on the notion that integrating aspects of EF allows children to control their behavior, remember instructions, pay attention, and complete learning tasks in school settings. In this study, we examined how well a measure of behavioral self-regulation tapped individual components of EF (cognitive flexibility, working memory, and inhibitory control) and how it predicted gains in academic achievement compared to these other EF measures.

THE HTKS MEASURE OF BEHAVIORAL SELF-REGULATION

The HTKS measure of behavioral self-regulation integrates aspects of EF into a short game appropriate for children aged 4–8 years. Using no materials but rather relying on interactions between the examiner and the child, the HTKS has three sections with up to four paired behavioral rules: "touch your head" and "touch your toes;" "touch your shoulders" and "touch your knees." Children first respond naturally, and then are instructed to switch rules by responding in the "opposite" way (e.g., touch their head when told to touch their toes). If children respond correctly after all four paired behavioral rules are introduced, the pairings are switched in the third section (i.e., head goes with knees and shoulders go with toes). In previous research (Cameron Ponitz et al., 2009; Wanless et al., 2011b; McClelland and Cameron, 2012), we have proposed that the HTKS measures behavioral self-regulation by requiring children to integrate into their behavior the following EF skills: (a) paying attention to the instructions, (b) using working memory to remember and execute new rules while processing the commands, (c) using inhibitory control through inhibiting their natural response to the test command while initiating the correct, unnatural response, and (d) using cognitive flexibility and working memory when rules accumulate and then change in the second and third sections.

Based on comparisons of HTKS scores to teacher ratings and parent reports of attention and inhibitory control, there is some evidence from previous research that the HTKS involves components of EF (McClelland et al., 2007a; Cameron Ponitz et al., 2009; Wanless et al., 2013). Other research has shown that the HTKS is significantly correlated with measures of working memory and requires children to successfully remember the changing rules of the task (Lan et al., 2011). However, some studies (including some

of our own previous work, e.g., Fuhs and Day, 2011; Lan et al., 2011; Turner et al., 2012) describe the task as predominately tapping inhibitory control or response inhibition. Thus, it is unclear if the HTKS is best aligned with one of the individual EF components, or if there is empirical evidence for it as a separate measure of behavioral self-regulation requiring the integration of multiple components. This issue has not been directly examined using multiple direct assessments of cognitive flexibility, working memory, and inhibitory control. Thus, a goal of the present study was to examine how the HTKS related to direct assessments of EF in a sample of children aged 3–7 years.

PREDICTORS OF ACADEMIC ACHIEVEMENT AND SCHOOL SUCCESS

Children's developmental trajectories are shaped by dynamic and interacting factors such as maturation, early experience, and brain development, especially in the prefrontal cortex (Diamond, 2002; Blair and Diamond, 2008; Blair and Raver, 2012). These factors also make the early childhood years a sensitive period for the development of behavioral self-regulation. Thus, given the potential malleability of behavioral self-regulation and related EF components, the early childhood years are an especially important time to examine relations between behavioral self-regulation and early academic achievement.

Of particular interest in the current study is the notion that behavioral self-regulation and EF processes are foundational for learning in a variety of domains especially in early childhood classrooms. Further, the pattern of skills that most strongly contributes to concurrent achievement may differ from skills that are important later in a child's developmental trajectory (Paris, 2005; Murrah, 2010). With regard to EF components, the development of inhibitory control is thought to occur first making it possible for children to demonstrate cognitive flexibility (Diamond et al., 2002; Best and Miller, 2010). These processes develop alongside working memory, though the development of this component is relatively more protracted with maturational improvements documented through adolescence (Best and Miller, 2010). One question these findings raise is which EF component(s) contribute the most to behavioral self-regulation at different ages across the early childhood span (and whether the components are the same or different across the prekindergarten and kindergarten years). In addition, the question of what skills and measures are most strongly associated with academic learning over the transition to school becomes important to address. This study examined the predictive validity of a measure of behavioral self-regulation and three EF component tasks to growth in academic achievement. We used random effect models and fixed effects models to examine predictive relations of each task to academic outcomes during the preschool and kindergarten years.

TESTING THE STRENGTH OF THE ASSOCIATION BETWEEN BEHAVIORAL SELF-REGULATION AND ACADEMIC OUTCOMES

A number of recent studies have examined the strength of associations between behavioral self-regulation and academic outcomes concurrently and longitudinally (Welsh et al., 2010; McClelland et al., 2013; Weiland and Yoshikawa, 2013). There is consistently strong evidence that behavioral self-regulation and EF significantly predict academic outcomes, even after controlling for baseline achievement levels, child IQ, and a host of demographic variables (e.g., McClelland et al., 2006, 2007a, 2013; Blair and Razza, 2007; Welsh et al., 2010; Moffitt et al., 2011). Relations have been especially strong for behavioral self-regulation and EF skills predicting growth in children's mathematics achievement (Blair and Razza, 2007; Cameron Ponitz et al., 2009; Bull et al., 2011).

Previous research on the relation between behavioral self-regulation, EF, and growth in academic outcomes has almost always utilized a *random effects approach (REA)*, in which the child is treated as a random draw from a distribution of individual differences in the rate of growth in academic skills. Such an approach can lead to biased estimates of how strongly a variable predicts growth when there are other time-invariant predictors of growth not included in the model (Clark and Linzer, 2012). An alternative approach, a *fixed effects approach (FEA)*, instead treats each child as a fixed effect (Allison, 2009), which eliminates this source of bias but at the expense of adding a large number of parameters associated with each child. The additional parameters (i.e., the fixed effect of each child in this case) mean the FEA can have lower power than the REA. To summarize, the REA can be used to examine inter-individual differences on behavioral self-regulation and explain these differences while modeling measured covariates that could be associated with behavioral and academic achievement (i.e., child IQ, age, parental education). The FEA can be used to investigate the association between intra-individual change over time in a child's behavioral self-regulation or EF skills and academic achievement.

In a study of 3- to 6-year-old children ($N = 794$), Willoughby and colleagues found that significant predictive relations between EF and academic achievement using a random effects approach became non-significant when using FEA (Willoughby et al., 2012b). Based on these results, Willoughby et al. (2012b) argued that the widely reported associations between EF and achievement might be spurious and driven by unmeasured time-invariant characteristics of the child. This argument, however, should be evaluated with caution. First, the null result could be attributable to a lack of power for a FEA to detect substantively significant effects rather than actual null effects. Second, the Willoughby et al. (2012b) study included just two time points (with an average of 4.4 months between time one and time two), so development in academic achievement may not have progressed sufficiently for individual differences in change to manifest. Furthermore, only two measures of EF (balance beam and pencil tapping) were used. Thus, it may not be surprising that there was no significant relation between the EF components that were measured and academic achievement in this study.

In addition, FEA findings tend to be sample specific (Allison, 2009; Clark and Linzer, 2012) making it difficult to generalize beyond any given study. This is partly because the sensitivity of a measure to change also depends on the validity and variability of the measure over time. This makes it important to replicate findings using different samples of children, with multiple measures and multiple time points. The current study sought to further test the strength of associations between behavioral self-regulation and academic achievement in young children using multiple measures of EF and behavioral self-regulation over the early school transition. Specifically, using both FEA and REA, we explored to

what extent four measures of EF and the HTKS measure of behavioral self-regulation significantly predicted achievement growth over four waves of data from the fall of prekindergarten to the spring of kindergarten. We anticipated that the two models would demonstrate the same overall pattern of results, especially for children's early mathematics skills. We anticipated that these results would be consistent across the two analytical approaches because we include more occasions of measurement and more measures of EF than the previous study using the lower powered FEA (Willoughby et al., 2012b).

THE PRESENT STUDY

The present study examined the longitudinal and psychometric properties of the HTKS measure of behavioral self-regulation by assessing: (1) construct validity through relations with traditional EF tasks, and (2) predictive validity for emergent literacy, vocabulary, and mathematics skills using random effects and fixed effects models. First, we anticipated that the HTKS would significantly relate to measures of cognitive flexibility, working memory, and inhibitory control based on previous research (McClelland et al., 2007a,b; Cameron Ponitz et al., 2009; Lan et al., 2011). Second, we considered predictive validity using random effects and fixed effects models between prekindergarten and kindergarten (over 4 time points). Based on previous research (e.g., Cameron Ponitz et al., 2009), we expected that compared to individual measures of cognitive flexibility, working memory, and inhibitory control, the HTKS would emerge as the strongest predictor of growth in academic achievement (literacy, vocabulary, and mathematics) in kindergarten. We also expected that the HTKS and measures of EF would be especially predictive of growth in early mathematics skills (Bull and Scerif, 2001; Cameron Ponitz et al., 2009; Bull et al., 2011).

METHOD

PARTICIPANTS AND PROCEDURE

The sample included 208 children (50% male) who participated in at least one wave of data collection (see **Table 1**). Families were recruited from 28 classrooms and 16 preschools located in the Pacific Northwest United States. The following kindergarten year, children were in 63 classrooms and 33 schools. Of the 208 children, 204 participated during wave 1; four children were not tested during wave 1 because they either refused testing sessions ($n = 3$) or parents asked for their child to be included during later waves ($n = 1$; see **Table 1** for total sample size by assessment and wave). Children and families were recruited through letters in an enrollment packet sent during the summer prior to the preschool year. Consent was obtained from a parent of all children in the study, and families were given $20 gift cards at each time point of the study.

Children were followed between preschool and kindergarten, with assessments in the fall and spring of each year (4 waves total). Children were assessed in English or Spanish in 2–3 sessions lasting 10–15 min each. About 50% of the children were enrolled in Head Start during the preschool year. At fall of preschool, children ranged in age from 36- to 65-months old ($M = 55.67$, $SD = 4.42$). Parent education ranged from about 5–23 years, with an average of approximately 3 years of college ($M = 14.80$, $SD = 3.68$ at baseline). Children were 61% White; 18% Latino; 0.5% African American; 1% Middle Eastern; 13% multiracial; and 1% other. About 14% of the sample was Spanish-speaking and were assessed in Spanish. In this sample, all Spanish-speaking children were identified as low-income. Moreover, low-income Spanish-speaking families reported significantly lower parent education levels, $[t_{(85)} = 4.958,\ p < 0.001]$, such that the parents of children who were Spanish-speaking reported lower levels of education ($M = 10.10$ years) than low-income English speakers ($M = 12.66$ years). In addition, compared to their low-income English-speaking peers, in the fall of preschool, Spanish-speaking children from low-income families scored significantly lower on the HTKS, $[t_{(95)} = 2.83,\ p = 0.006]$, some measures of EF [Dimensional Change Card Sort (DCCS), $t_{(99)} = 2.14$, $p = 0.035$, and Woodcock-Johnson Auditory Working Memory (WJ-WM) $t_{(98)} = 3.77$, $p < 0.001$], math, $[t_{(97)} = 4.41,\ p < 0.001]$, and literacy, $[t_{(97)} = 3.90,\ p < 0.001]$, but scored significantly higher on vocabulary scores, $[t_{(98)} = -2.51,\ p = 0.014]$.

Current research has focused on including diverse samples of children to appropriately assess EF in different populations. We included both Spanish-speaking and English-speaking children to examine our research questions in diverse groups. Previous research with different samples of low-income children who were Spanish-speaking or English-speaking did not find significant differences at the fall of prekindergarten in children's HTKS or EF scores (e.g., Wanless et al., 2011b; Schmitt et al., under review). Thus, we included both groups of children based on previous work evaluating the two groups separately.

MEASURES

Measures of behavioral self-regulation and EF

Children were assessed in preschool and kindergarten on the HTKS, Three-Dimensional Change Card Sort (DCCS), Day-Night Stroop task, the Auditory Working Memory subtest from the Woodcock-Johnson III Tests of Cognitive Abilities, and the Simon Says task. All tasks were counterbalanced to avoid order effects.

HTKS. The HTKS was used to assess children's behavioral self-regulation and requires cognitive flexibility, working memory, and inhibitory control (McClelland and Cameron, 2012). There are a total of 30 test items with scores of 0(*incorrect*), 1(*self-correct*), or 2(*correct*) for each item. A self-correct is defined as any motion to the incorrect response, but self-correcting and ending with the correct action. Scores range from 0 to 60 where higher scores indicate higher levels of behavioral self-regulation. The task takes approximately 5–7 min with strong inter-rater reliability ($\kappa = 0.90$; Cameron Ponitz et al., 2009; McClelland and Cameron, 2012). There are two parallel forms of the HTKS: A and B, which were given randomly in an alternating order of assessments over the four time points of the longitudinal study. Form A starts with head/toes and Form B starts with knees/shoulders. No significant differences have been found between the two versions of the task McClelland et al., 2007a; Cameron Ponitz et al., 2009; Wanless et al., 2011a; Bowles et al., submitted. The measure now incorporates three sections, the HTT (1 section of "opposites"),

Table 1 | Descriptive statistics of covariates, the HTKS and other EF measures, and achievement outcomes across four waves.

	Fall prekindergarten		Spring prekindergarten		Fall kindergarten		Spring kindergarten	
	N	*M* (*SD*)	*N*	*M* (*SD*)	*N*	*M* (*SD*)	*N*	*M* (*SD*)
Age	204	55.67 (4.42)	197	61.27 (4.45)	157	67.97 (3.88)	154	73.89 (3.87)
Percent male	204	50%	197	51%	157	54%	154	55%
Percent head start	204	50%	197	50%	157	45%[h]	154	47%[h]
Percent ELL[a]	204	14%	197	14%	157	13%	154	12%
Parent education	179	14.80 (3.68)	175	14.69 (3.67)	144	15.06 (3.76)	142	15.12 (3.76)
HTKS[b]	198	17.38 (16.92)	196	24.73 (18.61)	153	34.30 (17.60)	152	40.19 (15.27)
DCCS[c]	202	13.35 (6.72)	194	16.29 (6.09)	157	19.11 (4.49)	151	19.99 (3.47)
Day-Night Stroop	198	23.50 (9.12)	193	26.31 (7.90)	156	29.22 (4.29)	152	29.55 (3.95)
Working memory[d]	198	449.65 (15.10)	192	457.30 (18.69)	153	464.84 (19.75)	150	474.68 (19.29)
Simon Says	200	0.70 (1.38)	190	1.28 (1.92)	156	2.28 (1.93)	149	2.83 (1.84)
Mathematics[e]	197	409.31 (25.50)	194	419.99 (22.75)	155	434.14 (18.85)	152	444.39 (17.08)
Early literacy[f]	200	338.24 (25.65)	194	352.32 (26.45)	155	372.34 (29.46)	151	405.66 (36.93)
Vocabulary[g]	201	468.11 (14.06)	195	473.00 (11.72)	155	476.86 (12.21)	149	478.57 (11.34)

[a] *ELL = English Language Learner Status.*

[b] *The Head-Toes-Knees-Shoulders task.*

[c] *The Dimensional Change Card Sort task.*

[d] *The Woodcock-Johnson Auditory Working Memory subtest.*

[e] *The Woodcock-Johnson Applied Problems subtest.*

[f] *The Woodcock-Johnson Letter-Word Identification subtest.*

[g] *The Woodcock-Johnson Picture Vocabulary Subtest.*

[h] *Percent in Head Start is based on the child's prekindergarten year.*

the HTKS (2 sections, two sets of "opposites") and the HTKS—Extended (3 sections, adding a final rule switch). The task is available in a number of languages, is reliable, and significantly predicts academic outcomes in diverse samples (McClelland et al., 2007a,b; Wanless et al., 2011a; McClelland and Cameron, 2012; von Suchodoletz et al., 2013). Validity information for the current sample is presented in the Results below. Cronbach's alphas were computed in Mplus 7 using polychoric correlations, which are appropriate for categorical data. The HTKS in the current sample had Cronbach's alphas of 0.92, 0.94, 0.94, and 0.94 across the four waves of the study.

To assess inter-rater reliability in the current study, a random subsample of children ($n = 28$) was videotaped while being administered the HTKS task. Videotapes were later viewed and scored by an assessor who had not administered the original HTKS task to the child. We used double-coded HTKS sum scores analyzed with the default weighted kappa option in Stata (i.e., 1.00, 0.50, 0.00). The correlation between the double-coded HTKS scores was strong ($r = 0.88$, $p < 0.001$). Results showed high inter-rater agreement (92.29%), with a weighted Cohen's kappa of 0.79 ($p < 0.001$) indicating very strong inter-rater reliability for the HTKS task (Landis and Koch, 1977). To measure test-retest stability of the HTKS task in the current sample, Pearson's correlation coefficients for fall and spring HTKS scores were examined in prekindergarten and kindergarten (see **Table 2**). The average length of time between fall and spring HTKS task assessments was 5.64 months in prekindergarten ($SD = 0.57$, range = 4.17–7.16) and 5.84 months in kindergarten ($SD = 0.81$, range = 3.38–7.46). Results showed good test-retest stability with strong positive correlations between fall and spring HTKS total scores in both prekindergarten ($r = 0.60$, $p < 0.001$) and kindergarten ($r = 0.74$, $p < 0.001$).

Dimensional Change Card Sort (DCCS). Cognitive flexibility was assessed in English or Spanish using an adapted version of the Dimensional Change Card Sort (Deák, 2003; Hongwanishkul et al., 2005; Zelazo, 2006; Cepeda and Munakata, 2007), which is reliable and valid for children ages 3–5 years. Children were presented with cards that differed based on shape (i.e., dog, fish, bird), color (i.e., red, yellow, blue), and size (small, medium, large), and they were instructed to sort cards by each of the three dimensions. Children are first given six trials to sort by shape, then six trials to sort by color, then six trials to sort by size. If children scored at least five points on the sorting by size trial, children are given six more trials where they sorted cards by color and size depending on a border rule. The score is the sum of the total number of cards correctly sorted (1 = correct, 0 = incorrect) and scores can range from 0 to 24. In the current sample, the DCCS (using tetrachoric correlations) had Cronbach's alphas of 0.90, 0.92, 0.93, and 0.93 across four study waves.

Auditory working memory. The Auditory Working Memory test from the Woodcock-Johnson III Tests of Cognitive Abilities (Woodcock et al., 2001b) or The Bateria III Woodcock- Muñoz (Muñoz-Sandoval et al., 2005b) was used to assess children's working memory, the ability to remember and cognitively manipulate information. It demonstrates strong internal reliability: 0.93–0.96 for English-speaking preschool children and 0.77–0.79

Table 2 | Correlations of HTKS with other measures of EF during prekindergarten (*N* = 185–198) and kindergarten (*N* = 146–156).

Variable	1	2	3	4	5	6	7	8	9	10
1. Fall HTKS[a]	–	0.53***	0.29***	0.47***	0.44***	0.74***	0.33***	0.37***	0.59***	0.50***
2. Fall DCCS[b]	0.56***	–	0.23***	0.31***	0.41***	0.46***	0.51***	0.18*	0.41***	0.43***
3. Fall Day-Night Stroop	0.40***	0.36***	–	0.21**	0.13	0.22**	0.28***	0.63***	0.28***	0.12
4. Fall working memory[c]	0.41***	0.28***	0.20**	–	0.36***	0.38***	0.24**	0.19*	0.53***	0.25**
5. Fall Simon Says	0.38***	0.36***	0.31***	0.32***	–	0.42***	0.32***	0.19*	0.41***	0.58***
6. Spring HTKS[a]	0.60***	0.54***	0.34***	0.45***	0.32***	–	0.37***	0.27***	0.60***	0.48***
7. Spring DCCS[b]	0.46***	0.63***	0.31***	0.26***	0.28***	0.54***	–	0.25**	0.31***	0.36***
8. Spring Day-Night Stroop	0.31***	0.27***	0.41***	0.17*	0.13†	0.37***	0.32***	–	0.20*	0.24**
9. Spring working memory[c]	0.39***	0.40***	0.29***	0.38***	0.30***	0.39***	0.35***	0.26***	–	0.43***
10. Spring Simon Says	0.39***	0.47***	0.33***	0.40***	0.52***	0.54***	0.32***	0.21**	0.39***	–

Correlations on the bottom diagonal are for prekindergarten. Correlations on the top diagonal are for kindergarten.

[a] *The Head-Toes-Knees-Shoulders task.*

[b] *The Dimensional Change Card Sort task.*

[c] *The Woodcock-Johnson Auditory Working Memory subtest.*

†$p < 0.10$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

for Spanish-speaking children. Cronbach's alphas are not available for the current sample because scores were entered at the subtest level; however, it has a reported strong median split-half reliability of 0.93 for children 4–7 years old (Mather and Woodcock, 2001).

Day-Night Stroop task. Inhibitory control was assessed using the Day-Night Stroop task in English or Spanish (Gerstadt et al., 1994; Berwid et al., 2005). Children are shown a series of 16 cards with pictures of a sun or moon and asked to say the opposite of what they see, saying "day" for a moon and "night" for a sun. Each of the 16 items were coded as 0 for an incorrect response, 1 for a self-correct or similar (i.e., saying "sun" when the correct response is "day") response, or 2 for a correct response, with a possible range of 0–32. In the current sample, the Day-Night Stroop had Cronbach's alphas (using polychoric correlations) of 0.99, 0.99, 0.95, and 0.93 across four study waves.

Simon Says task. Inhibitory control was also assessed using Simon Says in English or Spanish. The measure is appropriate for prekindergarten and kindergarten children and has shown strong reliability and validity (Strommen, 1973; Carlson, 2005). Children are asked to perform an action only if the experimenter said "Simon says," but to remain still otherwise. Thus, the task measures inhibition but not inhibition plus activation, which is required for the HTKS. Of the 10 total trials, the 5 trials requiring inhibition are scored (0 = incorrect/imitation 1 = correct/anti-imitation) and children are given a proportion score of the number correct (anti-imitation) on these 5 trials. In the current sample, task scores ranged from 0 to 5 and had Cronbach's alphas (using tetrachoric correlations) of 0.95, 0.98, 0.93, and 0.91 across four waves.

We chose two measures of inhibitory control because we wanted to differentiate responses requiring inhibition only (children must stop or control motor activity), as in Simon Says, from those requiring inhibition of a dominant response plus activation of another, non-dominant response, as in Day-Night (Kochanska et al., 1996; Blair, 2003). This enabled us to examine which type of inhibition contributes the most to HTKS performance.

Academic achievement outcomes

Children's early reading, vocabulary, and math skills were assessed on the Woodcock Johnson Psycho-Educational Battery-III Tests of Achievement (WJ-III; Woodcock et al., 2001a) in English or the Batería III Woodcock-Muñoz (Muñoz-Sandoval et al., 2005a) in Spanish. Large-scale studies using item-response theory (IRT) have equated the English and Spanish WJ measures and indicate that they assess the same competencies (Woodcock and Muñoz-Sandoval, 1993). Recent research indicates no significant differences on scores between the English and Spanish versions of the WJ-III (Hindman et al., 2010).

Letter-word identification. Children's early literacy skills were measured using the Letter-Word Identification subtest of the WJ-III (Woodcock et al., 2001a) or The Bateria III Woodcock-Muñoz (Muñoz-Sandoval et al., 2005a). This test measures letter skills and developing word-decoding skills. Published split-half reliabilities for English-speaking preschool and kindergarten children range between 0.98–0.99 and 0.84–0.98 for Spanish-speaking children. The Letter-Word Identification subtest has a median split-half reliability of 0.98 for children 4–7 years old (Mather and Woodcock, 2001).

Picture vocabulary. Children's receptive and expressive vocabulary skills were assessed with the Picture Vocabulary subtest of the WJ-III or The Bateria III Woodcock-Muñoz. Published split-half reliabilities for English-speaking children range between 0.76–0.81 and 0.88–0.89 for Spanish-speaking children. The Picture Vocabulary subtest has a median split-half reliability of 0.73 for children 4–7 years old (McGrew and Woodcock, 2001).

Applied problems. The Applied Problems subtest of the WJ-III or The Bateria III Woodcock-Muñoz was used to assess children's early mathematical operations needed to solve practical problems. Published split-half reliabilities for 4- and 5-year-old English-speaking children are 0.92–0.94 and 0.93–0.95 for Spanish-speaking children. The Applied Problems subtest has a median split-half reliability of 0.92 for children 4–7 years old (McGrew and Woodcock, 2001).

Parent demographic questionnaires

All parents completed a demographic questionnaire including background characteristics such as child age, English Language Learner status, parent education level, and gender. These variables were used as covariates.

RESULTS

ANALYTIC STRATEGY

All research questions were addressed using Stata 13.1 (StataCorp, 2013). For construct validity, we first analyzed correlations between the HTKS and the four EF measures (the Day-Night Stroop, the DCCS, Simon Says, and the Woodcock-Johnson Working Memory subtest) for each wave. Then, we looked at multilevel models predicting HTKS scores with the four EF measures at each wave, controlling for child age, parent education, gender, Head Start status, and English Language Learner status. The *ICC*s for the HTKS across the four waves of data were: 0.12, 0.22, 0.15, and 0.10.

For predictive validity, we used multilevel models with generalized structural equation modeling in Stata 13.1, adjusting for the nested nature of the data (children within classrooms) and used a full information maximum likelihood estimator. For each random effects model, the models incorporated two waves of data, roughly 6 months apart during the same academic year (e.g., prekindergarten or kindergarten). In these models, the spring achievement variable was regressed on fall achievement, a single EF measure of interest, child age, parent education, gender, Head Start status, and English Language Learner status. The *ICC*s for the outcome achievement measures in the spring of prekindergarten (*ICC*s = 0.14–0.23) and kindergarten (*ICC*s = 0.22–0.27) suggested multilevel models were appropriate, and thus, all predictive models adjusted for this nesting.

Fixed effects analyses were estimated in Stata 13.1, with standard errors adjusted for clustering. In the fixed effects analyses, all four waves of data were analyzed simultaneously, such that all available data for each child from fall of prekindergarten to spring of kindergarten was modeled. In fixed effects analyses, associations of intra-individual change on predictors (i.e., EF) and outcomes (i.e., achievement) are of interest, thus no time-invariant covariates are included (as they were in the random effects model). Other than the effect of time, no time-varying covariates were used in these models (all time-invariant variables, measured and unmeasured, are incorporated in the estimate of the unit on the outcome).

Missing data, attrition, and descriptive statistics

Overall, there was relatively little missing data other than data lost due to attrition between the spring of prekindergarten and the fall of kindergarten (Waves 2–3). In the fall of prekindergarten (Wave 1), 204 children participated in the study. The most missing data on any assessment during the first wave occurred for the WJ-III Applied Problems subtest ($N = 197$) with 3.43% missing. In the spring of prekindergarten (Wave 2), a total of 197 children participated (97% retention from Wave 1 participants). The Simon Says task showed the most missing data with 3.55% missing.

In the fall of kindergarten (Wave 3, $N = 157$), 20.30% of the sample was lost due to attrition. Three covariates significantly predicted attrition from spring of prekindergarten to fall of kindergarten (year 1–2). Children were less likely to remain in the study if they were enrolled in Head Start during year 1, had parents with lower reported education levels, and were younger in age. Although differential attrition can lead to bias in parameter estimates, the use of covariates that predicted attrition (i.e., Head Start status, parental education, and age) with full information likelihood estimators are shown to provide reliable parameter estimates (Steiner et al., 2010).

In the fall of kindergarten (Wave 3), the task with the most missing was the HTKS with 2.55% missing data. From fall of kindergarten to spring of kindergarten (Wave 4, $N = 154$) there was a 98.09% retention rate. Of the participating children in Wave 4, the WJ-III Picture Vocabulary subtest and the Simon Says task showed the most missing with 3.25% missing data.

Descriptive statistics for covariates included in the models, parent-reported educational attainment, EF tasks, and achievement tasks are provided in **Table 1**. Furthermore, mean child performance improved in each EF measure and achievement measure across each wave of the study. In prekindergarten, children were clustered in 28 different classrooms ($M = 7.42$, range = 1–14), and by kindergarten, they had dispersed and were clustered in 63 different classrooms ($M = 2.50$, range = 1–10). We used full information maximum likelihood (FIML) to account for the small amount of missing data (Acock, 2012).

RQ 1: construct validity of the HTKS. Relations between the HTKS and each of the direct EF assessments of cognitive flexibility (DCCS), working memory (WJ-III Working Memory subtest), and inhibitory control (Day-Night, Simon Says) are presented for fall and spring of prekindergarten and kindergarten, with all correlations significant at = 0.001 (see **Table 2**). Overall, the HTKS was moderately correlated with the four direct assessments of EF throughout the four waves of data, suggesting convergent validity with traditional assessments of EF and construct validity that the HTKS assesses cognitive flexibility, working memory, and inhibitory control. For the fall of prekindergarten, the HTKS correlations with other EF tasks ranged from *r*s = 0.38–0.56 and for the spring of prekindergarten, correlations with other EF tasks ranged from *r*s = 0.37–0.54. For the fall of kindergarten, the HTKS correlations with other EF tasks ranged from *r*s = 0.29–0.53, and for the spring of kindergarten, correlations with other EF tasks ranged from *r*s = 0.27–0.60. Between prekindergarten and kindergarten, correlations among the EF measures ranged from *r*s = 0.20–0.56. The correlation between the HTKS and the DCCS was the strongest for the first three waves of data (*r*s from 0.46 to 0.56); however, by the spring of kindergarten

(wave four) the HTKS was slightly more related to the measure of working memory ($r = 0.60$; see **Table 2**).

After examining correlations, we used multilevel models treating the HTKS as an outcome predicted concurrently by the four EF measures and controlling for child age, parent education, gender, Head Start status, and English Language Learner status (see **Table 3**). Results were similar to the correlational findings but also revealed that (1) EF measures were independently related to the HTKS and (2) relative relations differed by wave. In the fall of prekindergarten, all four tasks significantly predicted the HTKS measure with the cognitive flexibility task (DCCS) having the relatively largest effect ($\beta = 0.36$, $p < 0.001$). In the spring of prekindergarten, the Simon Says inhibitory control task was the most predictive of HTKS scores ($\beta = 0.32$, $p < 0.001$), with only working memory being non-significant. In the fall of kindergarten, by contrast, the DCCS and working memory were the only significant predictors of the HTKS, with the DCCS having the largest effect ($\beta = 0.28$, $p < 0.001$). In the spring of kindergarten, the working memory and the Simon Says tasks were the only significant predictors, with working memory having the largest relative effect ($\beta = 0.42$, $p < 0.001$) on HTKS scores.

RQ 2: predictive validity of the HTKS and EF measures to academic outcomes. Random effects multilevel models were used to examine inter-individual differences on behavioral self-regulation and EF predicting improvement on achievement measures in each academic year (predictive validity). Results of multilevel regressions (i.e., predicting spring achievement from fall EF during the same academic year while controlling for fall achievement) indicated that Wave 1 prekindergarten performance on the HTKS, DCCS (cognitive flexibility), and Day-Night Stroop (inhibitory control) tasks predicted Wave 1-Wave 2 improvement in early mathematics ($\beta = 0.14$, $p = 0.007$; $\beta = 0.17$, $p = 0.002$; $\beta = 0.14$, $p = 0.006$, respectively; see **Table 4**). The DCCS and working memory tasks also predicted improvement in early vocabulary ($\beta = 0.11$, $p = 0.040$; $\beta = 0.10$, $p = 0.020$, respectively). None of the fall tasks significantly predicted early literacy improvement during the prekindergarten year.

Over the kindergarten year, Wave 3 scores on the HTKS, working memory, and Simon Says tasks predicted improvement in early mathematics ($\beta = 0.15$, $p = 0.018$; $\beta = 0.17$, $p = 0.002$; $\beta = 0.12$, $p = 0.038$, respectively; see **Table 4**). The HTKS task was the only task to significantly predict early literacy improvement ($\beta = 0.17$, $p = 0.001$). The HTKS, the Day-Night Stroop, and the Simon Says tasks significantly predicted kindergarten vocabulary improvement ($\beta = 0.16$, $p = 0.003$; $\beta = 0.10$, $p = 0.023$; $\beta = 0.14$, $p = 0.011$, respectively), with trend level effects on vocabulary for the DCCS ($\beta = 0.09$, $p = 0.095$).

Fixed effects models were run next to examine intra-individual change in behavioral self-regulation and EF predicting intra-individual change in the academic outcomes over the four time points. Results generally matched the findings of the random effects models, with some weaker associations: growth in the HTKS, the DCCS, and the Day-Night Stroop all significantly predicted growth in mathematics ($\beta = 0.10$, $p = 0.003$; $\beta = 0.09$, $p = 0.001$; $\beta = 0.07$, $p = 0.007$; respectively; see **Table 5**). For example, for each standard deviation increase on the HTKS, children made a 2.5 point gain on math. Thus, children who showed the most growth in behavioral self-regulation and EF also demonstrated the most growth in mathematics between prekindergarten and kindergarten. In addition, the Day-Night Stroop was the only task that significantly predicted improvement in vocabulary development ($\beta = 0.06$, $p = 0.039$). Thus, children making improvements in inhibitory control, as measured by the Day-Night Stroop task, also made significant improvements in vocabulary skills over the prekindergarten and kindergarten years. None of the measures significantly predicted growth in emergent literacy development between prekindergarten and kindergarten.

DISCUSSION

Results demonstrated that in prekindergarten and kindergarten, children who scored higher on the HTKS also performed better on each of the individual measures of EF (cognitive flexibility, working memory, and inhibitory control) although the strength of these relations varied over time. In addition, REA indicated the HTKS and the EF measures significantly predicted variation in

Table 3 | Construct validity: multilevel regressions of EF measures predicting HTKS during prekindergarten (*N* = 196–198) and Kindergarten (*N* = 152–153).

Predictor	Prekindergarten						Kindergarten					
	Fall HTKS[a]			Spring HTKS[a]			Fall HTKS[a]			Spring HTKS[a]		
	B	*SE*	β	*B*	*SE*	β	*B*	SE	β	*B*	*SE*	β
DCCS[b]	0.90	0.18	0.36***	0.72	0.18	0.24***	1.12	0.29	0.28***	0.42	0.30	0.10
Day-Night Stroop	0.27	0.13	0.14*	0.36	0.14	0.15*	0.41	0.26	0.10	0.13	0.25	0.03
Working memory[c]	0.21	0.07	0.19**	0.04	0.06	0.04	0.17	0.07	0.19*	0.34	0.06	0.42***
Simon Says	1.81	0.80	0.15*	3.12	0.57	0.32***	1.05	0.69	0.12	1.83	0.61	0.22**

Covariates (not shown) include parental education, child age (in months), Head Start status, gender, and English Language Learner status.

[a] *The Head-Toes-Knees-Shoulders task.*

[b] *The Dimensional Change Card Sort task.*

[c] *The Woodcock-Johnson Auditory Working Memory subtest.*

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 4 | Predictive validity: random effects models for the HTKS and other EF measures predicting achievement growth in prekindergarten (*N* = 194–195) and kindergarten (*N* = 149–152).

Predictor	Mathematics[d]			Early literacy[e]			Vocabulary[f]		
	B	*SE*	β	*B*	*SE*	β	*B*	*SE*	β
PREKINDERGARTEN									
HTKS[a]	0.19	0.07	0.14**	0.07	0.08	0.04	0.04	0.03	0.06
DCCS[b]	0.59	0.19	0.17**	0.23	0.21	0.06	0.20	0.10	0.11*
Day-Night Stroop	0.36	0.13	0.14**	0.09	0.15	0.03	0.09	0.07	0.07
Working memory[c]	0.05	0.07	0.03	0.09	0.08	0.05	0.08	0.03	0.10*
Simon Says	1.08	0.77	0.07	1.18	0.86	0.06	0.56	0.39	0.07
KINDERGARTEN									
HTKS[a]	0.15	0.06	0.15*	0.36	0.11	0.17**	0.10	0.04	0.16**
DCCS[b]	0.25	0.22	0.07	0.57	0.41	0.07	0.22	0.13	0.09†
Day-Night Stroop	0.01	0.19	0.00	0.50	0.40	0.06	0.27	0.12	0.10*
Working memory[c]	0.15	0.05	0.17**	0.14	0.10	0.07	0.02	0.03	0.04
Simon Says	1.03	0.50	0.12*	0.37	1.01	0.02	0.82	0.32	0.14*

Covariates (not shown) include parental education, child age (in months), Head Start status, gender, and English Language Learner status. Spring achievement gains control for fall achievement. Full Information Maximum Likelihood (FIML) estimation used to deal with missing data.

[a] The Head-Toes-Knees-Shoulders task.

[b] The Dimensional Change Card Sort task.

[c] The Woodcock-Johnson Auditory Working Memory subtest.

[d] The Woodcock-Johnson Applied Problems Subtest.

[e] The Woodcock-Johnson Letter-Word Identification subtest.

[f] The Woodcock-Johnson Picture Vocabulary Subtest.

*†p < 0.10; *p < 0.05; **p < 0.01.*

Table 5 | Predictive validity: fixed effects model coefficients for growth in HTKS and other EF measures predicting growth in achievement across four waves (*N* = 205–207).

Predictor	Mathematics[d]			Early literacy[e]			Vocabulary[f]		
	B	*SE*	β	*B*	*SE*	β	*B*	*SE*	β
HTKS[a]	0.13	0.04	0.10**	0.00	0.09	0.00	−0.00	0.02	−0.00
DCCS[b]	0.36	0.11	0.09**	−0.11	0.22	−0.02	0.09	0.06	0.04
Day-Night Stroop	0.26	0.09	0.08**	−0.04	0.14	−0.01	0.10	0.05	0.06*
Working memory[c]	−0.04	0.04	−0.03	0.06	0.08	0.03	−0.00	0.01	−0.01
Simon Says	−0.28	0.36	−0.02	−0.24	0.61	−0.01	−0.06	0.20	−0.01

[a] The Head-Toes-Knees-Shoulders task.

[b] The Dimensional Change Card Sort task.

[c] The Woodcock-Johnson Auditory Working Memory subtest.

[d] The Woodcock-Johnson Applied Problems Subtest.

[e] The Woodcock-Johnson Letter-Word Identification subtest.

[f] The Woodcock-Johnson Picture Vocabulary Subtest.

**p < 0.05; **p < 0.01.*

early achievement, with the strongest relations found for gains in early mathematics. In prekindergarten, measures of EF (especially the DCCS) were the strongest predictors of achievement in these models. In kindergarten, the HTKS was the most consistent predictor of achievement, although all measures of EF significantly predicted achievement depending on the time point. Results of the FEA found mostly consistent, albeit less strong, predictive relations compared to the random effects models.

CONSTRUCT VALIDITY OF THE HTKS

The current study sought to answer questions related to construct validity of a measure of behavioral self-regulation, called the HTKS. Previous research has differed on descriptions of what the HTKS measures, with some studies referring to the task as a measure of inhibitory control or response inhibition (Fuhs and Day, 2011; Lan et al., 2011), and some studies asserting evidence that it measures attention and working memory (McClelland

et al., 2007a; Cameron Ponitz et al., 2009; Lan et al., 2011). Adding to this complexity, we have conceptualized it theoretically as a measure of behavioral self-regulation, to recognize the social context in which the HTKS is administered and demonstrates validity. This is consistent with a recent distinction of *EF* as a top-down cognitive process, that enables the *self-regulation* of a more automatic, bottom-up set of processes, such as one would demonstrate in a spontaneous social setting like a classroom (Ursache et al., 2012). Nonetheless, little research has examined the HTKS alongside traditional EF component measures. Furthermore, scholars of behavioral self-regulation and EF have been criticized for producing a plethora of "conceptual clutter" and "measurement mayhem" in the conceptualization and measurement of these skills (Morrison and Grammer, in press). If the construct of behavioral self-regulation is important for children's short- and long-term academic achievement, equally important is understanding how tasks like the HTKS are related to measures of EF, including assessments of cognitive flexibility, working memory, and inhibitory control.

We also found that children who performed better on the HTKS had better cognitive flexibility, working memory, and inhibitory control in prekindergarten and kindergarten, though the strength of associations changed over time. At early time points, the HTKS was most related to cognitive flexibility (the DCCS) and inhibitory control (Simon Says, Day-Night Stroop). In contrast, at later time points, the HTKS was most strongly related to the measure of working memory, although it was still significantly correlated with the other measures of EF. Correlations and regressions suggest that the HTKS shares significant variance with all measures of EF in prekindergarten and kindergarten. However, and of particular note, the strength of these relations also varies over time as demonstrated in the correlations and the regression results. It is possible that these developmental differences in the patterns of performance may relate to underlying developmental trajectories. For example, more specific EF components such as cognitive flexibility or inhibitory control may be important for less complex tasks, while tasks capturing multiple EF components like the HTKS may be more important for more complex tasks later in development. It appears that the HTKS may tap different aspects of EF at different points in early childhood, although those conclusions are also limited by the EF measures themselves and the analyses, which do not allow us to explicitly compare parameter estimates. It is difficult to find a pure measure of working memory, inhibitory control, or cognitive flexibility, especially in young children. This has been termed "task impurity" in the literature and reflects the overlap of many EF components in early childhood (Landis and Koch, 1977; Hughes and Graham, 2002; Best et al., 2009).

In light of these caveats, the results of the present study lend support to previous research arguing that the HTKS taps multiple aspects of EF, and extends this research by suggesting that inhibitory control may predominate in determining HTKS performance for younger children, attentional or cognitive flexibility is relevant from ages 4 to 6 years, and working memory may contribute more to performance for older children (McClelland et al., 2007a; Cameron Ponitz et al., 2009; McClelland and Cameron, 2012). The result showing that the HTKS was most strongly related to the measure of working memory by the end of kindergarten is conceptually consistent with the task demands as children progress through the task. The second and third parts of the task require that children remember a newly introduced set of rules (Part II) and then switch those rules (Part III). This is supported by preliminary evidence showing adequate variability in the HTKS, especially the third part of the task through age eight (von Suchodoletz, in preparation).

PREDICTIVE VALIDITY OF THE HTKS AND EF MEASURES TO ACADEMIC OUTCOMES

We also examined the predictive validity of the HTKS and measures of EF using REA, which model inter-individual differences in behavioral self-regulation and EF on academic achievement; and FEA, which model intra-individual change in a child's behavioral self-regulation or EF skills and intra-individual change in academic achievement. In contrast to previous research that questioned the unique role of EF in achievement (e.g., Willoughby et al., 2012b), present results supported the predictive validity of both the HTKS and measures of EF to growth in academic achievement using a variety of analytic strategies. Results of both REA and FEA in this study supported previous research that links behavioral self-regulation and EF with achievement over the transition to formal schooling. Consistent with previous similar research treating the child as a random effect, each of the measures that we tested significantly predicted children's academic achievement gains in prekindergarten and kindergarten. Within the random effects framework, this pattern indicates that initial levels of behavioral self-regulation, cognitive flexibility, working memory, and inhibitory control are each foundational for learning over time (Blair and Razza, 2007; McClelland et al., 2007a; Blair and Diamond, 2008). Scholars have argued that such skills enable children to make sense of and manage the multiple demands of classroom settings, and help create a set of habits that lead to continued successes (Diamond, 2010; Blair and Raver, 2012). Results indicated that some of the EF measures (especially the DCCS) were the strongest predictors of achievement during the prekindergarten year, whereas the HTKS was the most consistent predictor of achievement in kindergarten. It is possible that individual measures of EF may be most predictive of earlier achievement, while the relative predictability of a behavioral self-regulation task for later achievement increases as children get older and are faced with more complex demands.

The finding that each of the individual measures, which were moderately correlated, were associated with achievement growth may indicate that the behaviors children need to learn are somewhat diverse or, at least, can be captured with multiple measures. At the same time, domain specificity was observed where, in general, measures of behavioral self-regulation and EF showed their strongest and most consistent relations with mathematics and vocabulary, as compared with literacy. The HTKS was also the only measure to significantly predict gains in literacy skills. Theoretically, we have argued that behavioral self-regulation requires that children integrate all aspects of EF and perform in ways that are especially relevant for learning in school settings; this position could be empirically confirmed if an integrative measure like the HTKS were the best predictor of learning

(McClelland and Cameron, 2012; McClelland et al., in press). The accumulating results for the HTKS using random effects models seem to support this position, but do not account for the fact that something else about the child, which both enables them to improve on the HTKS and to achieve academically over time, could explain the established links among the HTKS and later outcomes. Thus, we also examined our data using FEA.

Results of the FEA demonstrated similar, albeit less pronounced, patterns of predictability for the EF tasks and the HTKS measure of behavioral self-regulation. Measures of behavioral self-regulation (HTKS), cognitive flexibility (DCCS), and inhibitory control (Day-Night Stoop) significantly predicted growth in achievement between the fall of prekindergarten and the spring of kindergarten. The consistent significant finding for the HTKS and EF tasks and mathematics suggests that, during these early years, children who improved on measures of behavioral self-regulation and EF also demonstrated the most growth in mathematics. This finding matches a large body of evidence documenting strong links between children's EF and early mathematics (Blair and Razza, 2007; Bull and Lee, 2014). Reasons for this link can be tied to possible relations between specific components of EF and different aspects of early mathematics. For example, attentional shifting may be especially helpful for children to flexibly switch between multiple solutions to a math problem. In addition, inhibitory control may help children develop the types of learning-related behaviors that are needed to acquire early math skills, such as persistence and sequential problem-solving skills.

Our results suggest that aspects of EF and a measure of behavioral self-regulation are important for learning mathematics. Moreover, these results indicate that interventions to improve math might do well to target children's behavioral self-regulation as well as EF skills. Finally, children who made improvements on a measure of inhibitory control (the Day-Night Stroop task) also made significant gains in vocabulary skills between prekindergarten and kindergarten. Overall, this study, using two analytic methods, supports the robustness of the conclusion that behavioral self-regulation and EF component skills are important predictors of early academic achievement. However, in light of the reduced bias of unmeasured time-invariant variables, these results also suggest that the strength of prediction, although significant and substantial, may be somewhat lower than indicated by previous studies.

RESEARCH AND PRACTICAL IMPLICATIONS

At least two implications follow from the present study. First, the HTKS continues to demonstrate reliability and validity; and the measure seems to taps different aspects of EF although the strength of these relations varied over time between prekindergarten and the end of kindergarten. This is useful for researchers and practitioners who seek a short, economical, and psychometrically sound measure of behavioral self-regulation, which significantly predicts children's academic achievement—especially in mathematics—during the transition to formal schooling. Although researchers have emphasized the importance of using multiple measures of EF and behavioral self-regulation (Wiebe et al., 2008; Willoughby et al., 2012a), this may not always be feasible under time and budget constraints. The HTKS may be a practical alternative when it is not possible to use multiple measures and when predicting mathematics achievement is desirable (Duncan et al., 2007). Moreover, the minimal materials required for the task, coupled with its gross motor nature, make it an ecologically-appropriate measure for young children (McCabe et al., 2004).

The second implication is one for researchers, which points to continued examination of the constructs under investigation, but with the goals of parsimony, communication, and application. In early childhood, the dynamic development of multiple skill sets like EF and behavioral self-regulation means that, to some degree, we are studying a moving target. Furthermore, the use of distinct samples and measures introduces idiosyncrasies that contribute to the pattern of results for an individual study, yet are not well understood. It is one thing to draw conclusions about a construct from a single study, but researchers (including this author team) must also look across many studies to see the forest of EF components for the trees of what constructs and measures meaningfully predict whether or not children thrive in school. For example, the findings of this study may differ from those of Willoughby et al. (2012b) for multiple reasons, such as different measures or different sample characteristics.

It is also possible that relations between behavioral self-regulation and academic achievement may be reciprocal in young children. Recent research has demonstrated that an intervention focusing on academic skills in preschool led to significant improvements in academic outcomes and small improvements in EF (Weiland and Yoshikawa, 2013). Other research using cross-lagged models has found that the directionality is stronger from behavioral self-regulation to academic achievement than vice versa (Stipek et al., 2010), although more longitudinal work is needed. The overarching goal for scholars as well as teachers is not to increase scores on a behavioral self-regulation, EF, or achievement test *per se*, but to equip children with the general set of experiences and skills that will enable them to develop EF and demonstrate behavioral self-regulation within and beyond school settings (Blair and Raver, 2012). Furthermore, a number of interventions utilizing randomized controlled designs have demonstrated that interventions can significantly improve behavioral self-regulation and EF and academic achievement in young children (Bierman et al., 2008; Diamond and Lee, 2011; Raver et al., 2011; Tominey and McClelland, 2011; Schmitt et al., under review). Thus, despite continued refinement of terminology and methods, promoting behavioral self-regulation and EF in young children at home and at school is likely to help support their academic achievement and school success.

LIMITATIONS

This investigation had some limitations. First, although the sample was socioeconomically diverse (50% low-income), it was less ethnically diverse with 61% of the children being White. This concern is somewhat ameliorated by previous research indicating that the HTKS is associated with achievement in diverse groups of children from different cultures Wanless et al., 2011a,b; McClelland and Wanless, 2012; von Suchodoletz et al., 2013; Wanless et al., 2013. In addition, the sample in the current study represented the demographic characteristics of the region in

which it was drawn, but future research should include a greater diversity of children to better address this issue. Furthermore, covariates (i.e., Head Start status, parental education, and age) predicted attrition during year 1–2 of the study, and although these variables were used in the models with full information maximum likelihood to offset bias in estimates (Steiner et al., 2010), it is impossible to know if other unmeasured covariates were also related to attrition. Due to differential attrition and a non-random sample to begin with, generalizability of the findings might be limited and findings should be replicated in other studies. Second, it is possible that the presence of reduced variance (for instance, as seen in the Simon Says task at the fall of prekindergarten) could have limited the ability to detect significant associations between behavioral self-regulation and EF tasks and academic achievement outcomes. Third, although we used a variety of analytic strategies including FEA, we cannot infer causality from the results. As noted above, evidence from experimental studies indicate that improving children's behavioral self-regulation is likely to improve academic outcomes (Bierman et al., 2008; Diamond and Lee, 2011; Raver et al., 2011; Tominey and McClelland, 2011; Schmitt et al., under review), but more long-term research is needed. Finally, in the present study, all tasks were given to children by an assessor and not via computer. Thus, we were unable to measure information processing speed and use it as a control variable in our analyses. This is an avenue for future research.

CONCLUSIONS

We examined the construct validity of a measure of behavioral self-regulation, the HTKS, assessing associations with measures of EF including cognitive flexibility, working memory, and inhibitory control. A second aim examined predictive validity of growth in the HTKS and EF tasks to academic achievement growth between prekindergarten and the end of kindergarten. Results indicated that the HTKS taps aspects of cognitive flexibility, working memory, and inhibitory control, although the strength of these relations varied between prekindergarten and kindergarten. In addition, the HTKS and EF tasks significantly predicted growth in academic achievement over 2 years in both random effects and fixed effects analyses (FEA). These results indicate that the HTKS, which takes 5–7 min to administer and does not require extensive materials, may be a practical tool that predicts children's achievement over the transition to kindergarten.

ACKNOWLEDGMENTS

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A100566 to Oregon State University (M. McClelland, PI). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

REFERENCES

Acock, A. C. (2012). "What to do about missing values," in *Data Analysis and Research Publication. APA Handbook of Research Methods in Psychology*, Vol. 3, eds H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, and K. J. She (Washington, DC: American Psychological Association), 27–50.

Allison, P. D. (2009). *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage.

Barkley, B. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* 121, 65–94. doi: 10.1037/0033-2909.121.1.65

Baumeister, R. F., and Vohs, K. D. (2004). *Handbook of Self-Regulation: Research, Theory, and Applications*. New York, NY: Guilford.

Berwid, O. G., Curko Kera, E. A., Marks, D. J., Santra, A., Bender, H. A., and Halperin, J. M. (2005). Sustained attention and response inhibition in young children at risk for Attention Deficit/Hyperactivity Disorder. *J. Child Psychol. Psychiatry* 46, 1219–1229. doi: 10.1111/j.1469-7610.2005.00417.x

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., and Jones, L. L. (2009). Executive functions after age 5: changes and correlates. *Dev. Rev.* 29, 180–200. doi: 10.1016/j.dr.2009.05.002

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., and Domitrovich, C. E. (2008). Executive functions and school readiness intervention: impact, moderation, and mediation in the Head Start REDI program. *Dev. Psychopathol.* 20, 821–843. doi: 10.1017/S0954579408000394

Blair, C. (2003). Behavioral inhibition and behavioral activation in young children: relations with self-regulation and adaptation to preschool in children attending Head Start. *Dev. Psychobiol.* 42, 301–311. doi: 10.1002/dev.10103

Blair, C., and Diamond, A. (2008). Biological processes in prevention and intervention: the promotion of self-regulation as a means of preventing school failure. *Dev. Psychopathol.* 20, 899–911. doi: 10.1017/S0954579408000436

Blair, C., and Raver, C. C. (2012). Individual development and evolution: experiential canalization of self-regulation. *Dev. Psychol.* 48, 647–657. doi: 10.1037/a0026472

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Bronson, M. B. (2000). *Self-regulation in Early Childhood: Nature and Nurture*. New York, NY: The Guilford Press.

Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., and Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: sources of variation in emergent mathematic achievement. *Dev. Sci.* 14, 679–692. doi: 10.1111/j.1467-7687.2010.01012.x

Bull, R., and Lee, K. (2014). Executive functioning and mathematics achievement. *Child Dev. Perspect.* 8, 36–41. doi: 10.1111/cdep.12059

Bull, R., and Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Dev. Neuropsychol.* 19, 273–293. doi: 10.1207/S15326942DN1903_3

Cameron Ponitz, C. E., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., and Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Child. Res. Q.* 23, 141–158. doi: 10.1016/j.ecresq.2007.01.004

Cameron Ponitz, C. E., McClelland, M. M., Matthews, J. M., and Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Dev. Psychol.* 45, 605–619. doi: 10.1037/a0015365

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M., and Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333

Caughy, M. O., Mills, B., Owen, M. T., and Hurst, J. R. (2014). Emergent self regulation skills among very young ethnic minority children: a confirmatory factor model. *J. Exp. Child Psychol.* 116, 839–855. doi: 10.1016/j.jecp.2013.07.017

Cepeda, N., and Munakata, Y. (2007). Why do children perseverate when they seem to know better: graded working memory, or directed inhibition? *Psychon. Bull. Rev.* 14, 1058–1065. doi: 10.3758/BF03193091

Clark, T., and Linzer, D. (2012). Should I use fixed or random effects? *Working Paper.* Available onlie at: http://polmeth.wustl.edu/mediaDetail.php?docId=1315

Deák, G. O. (2003). "The development of cognitive flexibility and language abilities," in *Advances in Child Development and Behavior*, Vol. 31, ed R. Kail (San Diego: Academic Press), 271–327.

Diamond, A. (2002). "Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry," in

Principles of Frontal Lobe Function, eds D. T. Stuss and R. T. Knight (London: Oxford University Press), 466–503. doi: 10.1093/acprof:oso/9780195134971.003.0029

Diamond, A. (2010). The evidence base for improving school outcomes by addressing the whole child and by addressing skills and attitudes, not just content. *Early Educ. Dev.* 21, 780–793. doi: 10.1080/10409289.2010.514522

Diamond, A., Kirkham, N., and Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Dev. Psychol.* 38, 352–362. doi: 10.1037/0012-1649.38.3.352

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529

Dowsett, S. M., and Livesey, D. J. (2000). The development of inhibitory control in preschool children: effects of "Executive Skills" training. *Dev. Psychobiol.* 36, 161–174. doi: 10.1002/(SICI)1098-2302(200003)36:2<161::AID-DEV7>3.0.CO;2-0

Duckworth, A. L., Tsukayama, E., and May, H. (2010). Establishing causality using longitudinal hierarchical linear modeling: an illustration predicting achievement from self-control. *Soc. Psychol. Pers. Sci.* 1, 311–317. doi: 10.1177/1948550609359707

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P. et al. (2007). School readiness and later achievement. *Dev. Psychol.* 43, 1428–1446. doi: 10.1037/0012-1649.43.6.1428

Eisenberg, N., Duckworth, A. L., Spinrad, T. L., and Valiente, C. (2014). Conscientiousness: origins in childhood? *Dev. Psychol.* 50, 1331–1349. doi: 10.1037/a0030977

Fuhs, M. W., and Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at Head Start. *Dev. Psychol*. 47, 404–416. doi: 10.1037/a0021065

Gathercole, S. E., and Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *Br. J. Educ. Psychol.* 70, 177–194. doi: 10.1348/000709900158047

Geldhof, G. J., Little, T. D., and Colombo, J. (2010). "Self-regulation across the lifespan," in *The Handbook of Life-Span Development*, Vol. 2. *Social and Emotional Development*, eds M. E. Lamb, and A. Freund (Vol. eds R. M. Lerner) (Hoboken, NJ: Wiley), 116–157.

Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action: performance of children -7 years old on a stroop- like day-night test. *Cognition* 53, 129–153. doi: 10.1016/0010-0277(94)90068-X

Hindman, A. H., Skibbe, L. E., Miller, A., and Zimmerman, M. (2010). Ecological contexts and early learning: contributions of child, family, and classroom factors during Head Start, to literacy and mathematics growth through first grade. *Early Child. Res. Q.* 25, 235–250. doi: 10.1016/j.ecresq.2009.11.003

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., and Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: age-related changes and individual differences. *Dev. Neuropsychol.* 28, 617–644. doi: 10.1207/s15326942dn2802_4

Hughes, C., and Graham, A. (2002). Measuring executive functions in childhood: problems and solutions? *Child Adolesc. Ment. Health* 7, 131–142. doi: 10.1111/1475-3588.00024

Kail, R. V. (2003). "Information processing and memory," in *Crosscurrents in Contemporary Psychology*, eds M. H. Bornstein, L. Davidson, C. L. M. Keyes, and K. A. Moore (Mahwah, NJ: Lawrence Erlbaum), 269–279.

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., and Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Dev.* 67, 490–507. doi: 10.2307/1131828

Lan, X., Legare, C. H., Cameron Ponitz, C., Li, S., and Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: a cross-cultural analysis of Chinese and American preschoolers. *J. Exp. Child Psychol.* 108, 677–692. doi: 10.1016/j.jecp.2010.11.001

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. doi: 10.2307/2529310

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., and Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Dev. Psychol.* 46, 1062–1077. doi: 10.1037/a0020066

Mather, N., and Woodcock, R. W. (2001). *Examiner's Manual. Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

McCabe, L. A., Cunnington, M., and Brooks-Gunn, J. (2004). "The development of self-regulation in young children: individual characteristics and environmental contexts," in *Handbook of Self-regulation: Research, Theory, and Applications*, eds R. F. Baumeister and K. D. Vohs (New York, NY: Guilford Press), 340–356.

McClelland, M. M., Acock, A. C., and Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Child. Res. Q.* 21, 471–490. doi: 10.1016/j.ecresq.2006.09.003

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., and Stallings, M. C. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childh. Res. Q.* 28, 314–324. doi: 10.1016/j.ecresq.2012.07.008

McClelland, M. M., and Cameron, C. E. (2012). Self-Regulation in early childhood: improving conceptual clarity and developing ecologically valid measures. *Child Dev. Perspect.* 6, 136–142. doi: 10.1111/j.1750-8606.2011.00191.x

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., and Morrison, F. J. (2007a). Links between behavioral regulation and preschoolers' literacy, vocabulary and math skills. *Dev. Psychol.* 43, 947–959. doi: 10.1037/0012-1649.43.4.947

McClelland, M. M., Cameron, C. E., Wanless, S. B., and Murray, A. (2007b). "Executive function, behavioral self-regulation, and social-emotional competence: links to school readiness," in *Contemporary Perspectives on Social Learning in Early Childhood Education*, eds O. N. Saracho and B. Spodek (Charlotte, NC: Information Age), 83–107.

McClelland, M. M., Cameron Ponitz, C., Messersmith, E., and Tominey, S. (2010). "Self-regulation: the integration of cognition and emotion," in *Handbook of Life-Span Human Development,* Vol. 1. *Cognition, Biology And Methods*, ed W. F. Overton (Editor-in-chief: R. M. Lerner) (Hoboken, NJ: Wiley and Sons), 509–553.

McClelland, M. M., Geldhof, G. J., Cameron, C. E., and Wanless, S. B. (in press). "Development and self-regulation," in *Theory and Method*. Vol. 1. *Handbook of Child Psychology and Developmental Science,* 7th Edn, eds W. F. Overton and P. C. M. Molenaar (Editor-in-Chief: Richard M. Lerner) (Hoboken, NJ: Wiley).

McClelland, M. M., and Wanless, S. B. (2012). Growing up with assets and risks: the importance of self-regulation for academic achievement. *Res. Hum. Dev.* 9, 278–297. doi: 10.1080/15427609.2012.729907

McGrew, K. S., and Woodcock, R. W. (2001). *Technical Manual. Woodcock-Johnson III*. Itasca, IL: Riverside Publishing.

Miyake, A. U., Friedman, N. P., Emerson, M. J., Witzki, A. H., and Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H. et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108

Morrison, F. J., and Grammer, J. K. (in press). "Conceptual clutter and measurement nayhem: proposals for cross disciplinary integration in conceptualizing and measuring executive function," in *Executive Function in Preschool Age Children: Integrating Measurement, Neurodevelopment and Translational Research*, eds J. Griffin, P. McCardle, and L. Freund (Washington, DC: American Psychological Association).

Morrison, F. J., Ponitz, C. C., and McClelland, M. M. (2010). "Self-regulation and academic achievement in the transition to school," in *Child development at the Intersection of Emotion and Cognition*, eds S. D. Calkins and M. Bell (Washington, DC: American Psychological Association), 203–224.

Müller, U., Dick, A. S., Gela, K., Overton, W. F., and Zelazo, P. D. (2006). The role of negative priming in preschoolers' flexible rule use on the Dimensional Change Card Sort Task. *Child Dev.* 77, 395–412. doi: 10.1111/j.1467-8624.2006.00878.x

Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K. S., and Mather, N. (2005a). *Bateria III Woodcock- Muñoz: Pruebas de Aprovechamiento*. Itasca, IL: Riverside Publishing.

Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K. S., and Mather, N. (2005b). *Bateria III Woodcock-Muñoz: Pruebas de Habilidades Cognitivas*. Itasca, IL: Riverside Publishing.

Murrah, W. M. 3rd. (2010). *Comparing Self-Regulatory and Early Academic Skills as Predictors of Later Math, Reading, and Science Elementary School Achievement.* 72 dissertation/thesis, ProQuest Information and Learning.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Read. Res. Q.* 40, 184–202. doi: 10.1598/RRQ.40.2.3

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., and Pressler, K. (2011). CSRP's impact on low-income preschoolers' preacademic skills: self-regulation as a mediating mechanism. *Child Dev.* 82, 362–378. doi: 10.1111/j.1467-8624.2010.01561.x

Rennie, D. A. C., Bull, R., and Diamond, A. (2004). Executive functioning in preschoolers: reducing the inhibitory demands of the dimensional change card sort task. *Dev. Neuropsychol.* 26, 423–443. doi: 10.1207/s15326942dn2601_4

Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., and Brock, L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Dev. Psychol.* 45, 958–972. doi: 10.1037/a0015861

Rothbart, M. K., and Posner, M. I. (2005). Genes and experience in the development of executive attention and effortful control. *New Dir. Child Adolesc. Dev.* 109, 101–108. doi: 10.1002/cd.142

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., and Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14931–14936. doi: 10.1073/pnas.0506897102

StataCorp. (2013). *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.

Steiner, P. M., Cook, T. D., Shadish, W. R., and Clark, M. H. (2010). The importance of covariate selection in controlling for bias in observational studies. *Psychol.Methods* 15, 250–267. doi: 10.1037/a0018719

Stipek, D., Newton, S., and Chudgar, A. (2010). Learning-related behaviors and literacy achievement in elementary school-aged children. *Early Child. Res. Q.* 25, 385–395. doi: 10.1016/j.ecresq.2009.12.001

Strommen, E. A. (1973). Verbal self-regulation in a children's game: impulsive errors on "Simon says". *Child Dev.* 44, 849–853. doi: 10.2307/1127737

Tominey, S. L., and McClelland, M. M. (2011). Red light, purple light: findings from a randomized trial using circle time games to improve behavioral self-regulation in preschool. *Early Educ. Dev.* 22, 489–519. doi: 10.1080/10409289.2011.574258

Turner, K. A., Lipsey, M. W., Farran, D. C., Dong, N., Wilson, S. J., Fuhs, M. W. et al. (2012). "Academically relevant measures of self-regulation: Development and validation of assessments for preschool children," in *Paper Presented at the Society for Research on Educational Effectiveness* (Washington, DC).

Ursache, A., Blair, C., and Raver, C. C. (2012). The promotion of self-regulation as a means of enhancing school readiness and early achievement in children at risk for school failure. *Child Dev. Perspect.* 6, 122–128. doi: 10.1111/j.1750-8606.2011.00209.x

von Suchodoletz, A., Gestsdottir, S., Wanless, S. B., McClelland, M. M., Birgisdottir, F., Gunzenhauser, C. et al. (2013). Behavioral self-regulation and relations to emergent academic skills among children in Germany and Iceland. *Early Childh. Res. Q.* 28, 62–73. doi: 10.1016/j.ecresq.2012.05.003

von Suchodoletz, A., Trommsdorff, G., Heikamp, T., Wieber, F., and Gollwitzer, P. M. (2009). Transition to school: the role of kindergarten children's behavior regulation. *Learn. Individ. Differ.* 19, 561–566. doi: 10.1016/j.lindif.2009.07.006

Wanless, S. B., McClelland, M. M., Acock, A. C., Cameron Ponitz, C., Son, S. H., Lan, X. et al. (2011a). Measuring behavioral regulation in four societies. *Psychol. Assess.* 23, 364–378. doi: 10.1037/a0021768

Wanless, S. B., McClelland, M. M., Lan, X., Son, S. H., Cameron, C. E., Morrison, F. J. et al. (2013). Gender differences in behavioral regulation in four societies: the United States, Taiwan, South Korea, and China. *Early Childh. Res. Q.* 28, 621–633. doi: 10.1016/j.ecresq.2013.04.002

Wanless, S. B., McClelland, M. M., Tominey, S. L., and Acock, A. C. (2011b). The influence of demographic risk factors on children's behavioral regulation in prekindergarten and kindergarten. *Early Educ. Dev.* 22, 461–488. doi: 10.1080/10409289.2011.536132

Weiland, C., and Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Dev.* 84, 2112–2130. doi: 10.1111/cdev.12099

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., and Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *J. Educ. Psychol.* 102, 43–53. doi: 10.1037/a0016738

Wiebe, S. A., Espy, K. A., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575

Willoughby, M. T., Blair, C. B., Wirth, R. J., and Greenberg, M. (2012a). The measurement of executive function at age 5: psychometric properties and relationship to academic achievement. *Psychol. Assess.* 24, 226–239. doi: 10.1037/a0025361

Willoughby, M. T., Kupersmidt, J. B., and Voegler-Lee, M. E. (2012b). Is preschool executive function causally related to academic achievement? *Child Neuropsychol.* 18, 79–91. doi: 10.1080/09297049.2011.578572

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001a). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing Company.

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001b). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Woodcock, R. W., and Muñoz-Sandoval, A. F. (1993). An IRT approach to cross-language test equating and interpretation. *Eur. J. Psychol. Meas.* 32, 233–241.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): a method for assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Zelazo, P. D., and Cunningham, W. A. (2007). "Executive function: mechanisms underlying emotion regulation," in *Handbook of Emotion Regulation*, ed J. Gross (New York, NY: Guilford Press), 135–158.

Zelazo, P. D., and Müller, U. (2002). "Executive function in typical and atypical development," in *Blackwell Handbook of Childhood Cognitive Development*, ed U. Goswami (Malden, MA: Blackwell Publishing), 445–469.

Zimmerman, B. J. (2000). "Attaining self-regulation: a social cognitive perspective," in *Handbook of Self-regulation*, eds M. Boekaerts, P. R., Pintrich, and M. Ziedner (San Diego: CA: Academic Press), 13–39.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 January 2014; accepted: 28 May 2014; published online: 17 June 2014.
*Citation: McClelland MM, Cameron CE, Duncan R, Bowles RP, Acock AC, Miao A and Pratt ME (2014) Predictors of early growth in academic achievement: the head-toes-knees-shoulders task. Front. Psychol. **5**:599. doi: 10.3389/fpsyg.2014.00599*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 McClelland, Cameron, Duncan, Bowles, Acock, Miao and Pratt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 April 2014
doi: 10.3389/fpsyg.2014.00287

Sorting Test, Tower Test, and BRIEF-SR do not predict school performance of healthy adolescents in preuniversity education

Annemarie Boschloo[1,2]*, Lydia Krabbendam[1], Aukje Aben[3], Renate de Groot[4] and Jelle Jolles[1]

[1] *Department of Educational Neuroscience, Faculty of Psychology and Education, LEARN! Research Institute, VU University Amsterdam, Amsterdam, Netherlands*
[2] *Hogeschool iPabo, Alkmaar, Netherlands*
[3] *Het Bouwens van der Boijecollege, Panningen, Netherlands*
[4] *Centre for Learning Sciences and Technologies, Open Universiteit Nederland, Heerlen, Netherlands*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
John Best, University of British Columbia, Canada
Robert D. Latzman, Georgia State University, USA

****Correspondence:***
Annemarie Boschloo, Department of Educational Neuroscience, Faculty of Psychology and Education, LEARN! Research Institute for Learning and Education, VU University Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, Netherlands
e-mail: a.m.boschloo@vu.nl

Executive functions (EF) such as self-monitoring, planning, and organizing are known to develop through childhood and adolescence. They are of potential importance for learning and school performance. Earlier research into the relation between EF and school performance did not provide clear results possibly because confounding factors such as educational track, boy-girl differences, and parental education were not taken into account. The present study therefore investigated the relation between executive function tests and school performance in a highly controlled sample of 173 healthy adolescents aged 12–18. Only students in the pre-university educational track were used and the performance of boys was compared to that of girls. Results showed that there was no relation between the report marks obtained and the performance on executive function tests, notably the Sorting Test and the Tower Test of the Delis-Kaplan Executive Functions System (D-KEFS). Likewise, no relation was found between the report marks and the scores on the Behavior Rating Inventory of Executive Function—Self-Report Version (BRIEF-SR) after these were controlled for grade, sex, and level of parental education. The findings indicate that executive functioning as measured with widely used instruments such as the BRIEF-SR does not predict school performance of adolescents in preuniversity education any better than a student's grade, sex, and level of parental education.

Keywords: neuropsychology, executive functions, adolescence, academic performance, education

INTRODUCTION

At school, adolescents often get complex assignments and have to do homework for various courses simultaneously. In addition, they have to decide which combinations of courses to follow, which in turn may affect their possibilities for higher education and future careers. Therefore, the adolescent student needs to develop higher cognitive skills, such as self-monitoring, planning and organizing, in order to perform well. It is unclear, however, whether the development of these functions also predicts adolescents' school performance. Insight into the cognitive predictors of school performance is relevant for school (neuro) psychologists and other professionals who work with adolescents. They often have to estimate how scores on intelligence tests and neuropsychological tests are related to task performance in adolescents' daily life, for example to performance at school.

The neuropsychological measures often used for estimating performance in daily life are executive function tests (Gioia and Isquith, 2004; Chan et al., 2008). Executive functions are the functions necessary for goal-directed behavior (e.g., Best and Miller, 2010). A wide range of executive functions have been described in literature, such as inhibition, updating working memory, shifting, planning, organization skills, attentional control, and self-control (Alvarez and Emory, 2006; Best and Miller, 2010; Hofmann et al., 2012). However, concerns have been raised about the *ecological validity* of executive function tests; that is, how well they predict performance in daily life (Gioia and Isquith, 2004; Chan et al., 2008; Olson et al., 2013). Previous studies that related executive function tests to school performance in adolescents found mixed results (e.g., Gioia and Isquith, 2004; St Clair-Thompson and Gathercole, 2006; Chan et al., 2008; Latzman et al., 2010; Best et al., 2011), which we will address in depth below. These mixed results may have been caused by a lack of control for important confounders (Willoughby et al., 2012). Therefore, the present study set out to investigate whether performance on executive function tests is related to school performance in a highly controlled sample of adolescents.

Neuroscientific research associates executive functions with the functioning of neural networks between several brain areas including, but not restricted to, the prefrontal brain areas (Alvarez and Emory, 2006). These brain areas develop during childhood through adolescence until early adulthood (Gogtay et al., 2004; Giedd, 2008). Neuropsychological studies have confirmed that executive functions develop during this time period, with some functions becoming fully developed earlier than others (Anderson, 2002; Best and Miller, 2010). Considering the nature of executive functions, and the fact that they are still developing

during adolescence, it is likely that adolescents' school performance is related to the degree of maturation of relevant executive functions.

As we mentioned earlier, studies investigating the relation between executive functions and school performance have reported mixed results. These studies can be classified by the outcome measures they used to assess school performance. School performance can be measured with various outcome measures, such as report marks or performance on standardized tests. Of these two measures of school performance, report marks have the highest ecological validity, since they are relevant for students' daily lives. Decisions on passing or failing a course or grade are made based on report marks.

Only a few studies have investigated the relation between executive functions and report marks. Most of these studies were conducted with young adolescents, aged 12–13 (Veenman et al., 2005; Checa et al., 2008; Checa and Rueda, 2011). Results showed that executive functions such as executive attention (Checa et al., 2008), and metacognitive (executive) skills (Veenman et al., 2005), partially predicted report marks for mathematics. Attention and effortful control were found to be related not only to performance in mathematics, but also to the average report mark of all subjects at the end of the academic year (Checa and Rueda, 2011). In primary school children in third grade, executive function tests such as the Trail Making Test and the Tower of Hanoi did not relate to report marks. However, performance on a classroom-based planning task and teacher reports on children's time management skills were related to report marks (Cohen et al., 1995).

Most studies that have investigated the relation between executive functions and school performance in adolescents did not consider report marks, but looked at the outcomes of standardized performance tests. Standardized tests are equal for all students, and scores are not dependent on a student's school, class or teacher, as is the case with report marks. Studies on the relation between performance on standardized tests and executive functions showed that girls' performance on mathematics in adolescence and early adulthood was predicted by executive functioning measured in childhood, and especially by the score obtained on the Rey Osterrieth Complex Figure (Miller and Hinshaw, 2010; Miller et al., 2012). Furthermore, in a cross-sectional study, three complex executive function measures from the Cognitive Assessment System were related to school performance on reading and mathematics in children and adolescents aged 5–17 (Best et al., 2011). Other studies found that not all executive function tests contributed equally to various academic skills. Results from one study on adolescent boys aged 11–16 (Latzman et al., 2010) showed that: the Delis-Kaplan Executive Functions System (D-KEFS) composite score for conceptual flexibility was related to performance in reading and science; the monitoring composite was related to reading and social studies; and inhibition was related to mathematics and science. Another study on adolescents aged 11–12 (St Clair-Thompson and Gathercole, 2006) reported that: updating was related to performance in English and mathematics; inhibition was related to English, mathematics, and science; and that shifting was not related to school performance.

As these studies show, performance on some executive function tests appears to be related to school performance. However, the study results vary when it comes to determining which specific executive functions are related to different school subjects, and they are inconclusive about the exact extent of the relationships. Studies that report high correlations between executive functions and school performance (between 0.30 and 0.50 or higher), often used a sample that had diverse socioeconomic backgrounds, or they did not control for sex or intelligence (e.g., St Clair-Thompson and Gathercole, 2006; Best et al., 2011). Studies that did control for these factors generally found lower correlations (around 0.10–0.25) (Latzman et al., 2010; Miller and Hinshaw, 2010). Clearly, in order to investigate the association between executive functions and school performance, it is crucial to carefully control for confounders. In addition, research in younger children shows that the relationship between executive functions and school performance is confounded by unmeasured variables that are constant over time, such as household or care-giver characteristics (Willoughby et al., 2012).

As it is impossible to measure all potential confounders that may affect executive functions, the current study used a homogeneous sample. In that way, we were able to control for many known and unknown variables. Our sample consisted of students in the preuniversity educational track, which is the most advanced track of the Dutch secondary school system; the top 20% of all students in Dutch secondary education are in this track (Ministry of Education, Culture and Science, 2009). Therefore, all our participants were high-performing students. Moreover, we selected students who had never repeated or skipped a grade in school. Studies show that students who have repeated or skipped a grade have different profiles with regard to a range of school related variables compared to students with a regular educational career (Jimerson, 2001; Steenbergen-Hu and Moon, 2011). In addition, given that the former are a year older or younger than their classmates, they are most probably at a different stage of biological and psychological development. Finally, we reduced the effects of medical factors that may influence the relation between executive functions and school performance, such as past brain trauma, a developmental disorder, or medication use, by including only healthy, normally developing adolescents. Because of these selection criteria, our sample was homogeneous with regard to both ability level and developmental history.

The current study investigated one possible confounder in particular, namely sex. Sex is well known for its influence on school performance, as girls and boys excel at different subjects (Machin, 2005; Van Langen et al., 2006; Clark et al., 2008; Driessen and Van Langen, 2010). There is also growing evidence to support the conclusion that the neuropsychological performance of boys and girls differs in the school setting (Martens et al., 2011; Dekker et al., 2013a). In addition, a recent study has reported differences in executive functioning related to school performance in adolescents of different sex (Coenen et al., 2011). Other studies have shown that girls perform better in the school setting because they are better at self-control and self-discipline (Downey et al., 2005; Duckworth and Seligman, 2006; Hyde et al., 2007; Steinmayr and Spinath, 2008), which is interesting as there are indications that executive functions subserve self-control and self-discipline

(Hofmann et al., 2012). Finally, evidence is accumulating that boys and girls differ in brain maturation, especially during an extended period in adolescence, with boys lagging behind (Giedd, 2008; Lenroot and Giedd, 2010). This suggests that biological factors pertaining to brain development may underly the complex relation between executive functions and school performance. Other factors may influence the relation between sex and school performance as well. For example, boys are more likely to show work-avoidant motivational strategies than girls in secondary education (Dekker et al., 2013b). It may thus be that a relation between executive function and school performance is visible in girls, but not in boys. This hypothesis is addressed in the present study.

In sum, the main aim of the current study was to investigate the respective relations between three different measures of executive functions and school performance, while keeping close control of confounders. In addition, we investigated whether these relations were moderated by sex. We investigated a homogeneous sample of 173 healthy adolescents, all secondary school students in the pre-university educational track. Two objective, performance-based neuropsychological tests were used to measure categorizing and shifting (Sorting Test from the D-KEFS) and planning skills (Tower Test from the D-KEFS). These tests are suitable for administration to adolescents, and measure executive functions that are still developing at this age range (Delis et al., 2001; Huizinga et al., 2006; Luciana et al., 2009). Furthermore, we also administered the Behavior Rating Inventory of Executive Function—Self-Report Version (BRIEF-SR) (Guy et al., 2004). This questionnaire has been developed to measure a wide range of executive functions based on their appearance in real-world behavior. Therefore, the BRIEF-SR has been claimed to be a more ecologically valid measure of executive functions than objective neuropsychological tests (Gioia and Isquith, 2004; Guy et al., 2004; Olson et al., 2013). We used report marks to measure school performance, since these are most relevant to adolescents themselves. School performance was measured with the end-of-term report marks for Dutch (the native language), English as a foreign language, and mathematics. Based upon the assumption that the BRIEF-SR is more ecologically valid than objective executive function tests, we hypothesized that the BRIEF-SR would predict report marks better than the objective tests. In addition, we hypothesized that the relation between executive functions and school performance would be different for boys and girls.

METHODS

PARTICIPANTS

Participants came from seven secondary schools in the south of the Netherlands. They were in grade 7, 9, or 11 of the pre-university educational track. This is the most advanced track in Dutch secondary education; the top 20% of all students in Dutch secondary education are in this track (Ministry of Education, Culture and Science, 2009). Participants had not repeated or skipped a grade. Furthermore, participants had the Dutch nationality, had no learning disorders, psychiatric disorders or developmental disorders, did not use medication that influences cognitive functions and did not have a history of brain damage with a loss of consciousness of more than 30 min. These criteria were measured with a questionnaire that was completed by the parents.

The participants themselves and the parents of under-aged participants had to give permission for participation. Participants received a monetary reward for participation. The Ethical Committee of the Faculty of Psychology of Maastricht University approved the research protocol.

MEASURES

Executive functions

Objective measures of executive functions were acquired with the Sorting Test and the Tower Test from the Delis-Kaplan Executive Functions System (D-KEFS) (Delis et al., 2001). The Sorting Test is a card sorting test that aims to measure categorization skills and set shifting. No Dutch version existed; therefore we translated the words on the cards, and changed some words to make all original sorts possible. The free sorting condition was used. Outcome measure of the Sorting Test was the number of confirmed correct sorts (range: 0–16). The Tower Test measures planning, and has a strong learning component due to the nature of the items. The raw total achievement score was used as outcome measure (range: 0–30).

As a subjective measure of executive functions, the Behavior Rating Inventory of Executive Function—Self Report Version (BRIEF-SR) (Guy et al., 2004) was used. The BRIEF-SR is an 80-item questionnaire, especially developed for adolescents, in which they have to indicate how often the described behaviors had been a problem in the past six months (never, sometimes or often). The items can be grouped into 8 scales that measure the following executive functions: Inhibit, Shift, Emotional Control, Monitor (together: the Behavioral Regulation Index, BRI), and Working Memory, Plan/Organize, Organization of Materials, and Task Completion (together: the Metacognition Index, MCI). A higher score on the MCI and the BRI indicates more problems with executive functioning. Following the official Dutch translation of the BRIEF Parent Version (by Smidts and Sergeant), the BRIEF-SR was translated into Dutch. Few items of the BRIEF-SR are different from those in the Parent Version. These were translated by a native English-Dutch bilingual psychologist, and reviewed by another psychologist. The internal consistency of this Dutch version of the BRIEF-SR was $r = 0.89$ for the BRI and $r = 0.91$ for the MCI.

Report marks

End of term report marks (ranging from 1.0 = very bad to 10.0 = outstanding) for Dutch, English, and mathematics were acquired through the schools' administration. Dutch, English, and mathematics are the first three main goals of secondary education in the Netherlands (Ministry of Education, Culture and Science, 2006) and are valid estimators of school performance (Reed et al., 2010). These report marks are the result of multiple smaller and larger tests and assessments (at least more than 4) that were administered during one school year. The tests and assessments were part of the teaching method or were developed by teachers themselves, and could consist of various assessment methods, e.g., paper-and-pencil tests, essays, presentations. Because the schools in the sample used different grading policies, each school's report marks

were transformed into z-scores, based on the school's mean report mark and its standard deviation. In this way, the distribution of scores was similar for each school.

Demographics

Participants reported age and sex. Parents reported both parents' education level. Level of parental education (LPE) was defined as the highest education of the two. LPE was medium when the parents had junior vocational or junior general secondary education and high when they had senior vocational or academic education.

PROCEDURE

Adolescents were recruited through letters that were distributed at the seven schools by the researchers. All students were in grade 7, 9, or 11 at the start of the study. Because the study started at the end of a school year, 50.9% of the adolescents were tested in the new school year, and were therefore in grade 8, 10, or 12 when they participated. A trained psychologist administered tests and questionnaires in a quiet room at school. Administration took approximately 1.5 h. Adolescents participated during school time and therefore missed certain lessons.

ANALYSES

All analyses were performed with SPSS Statistics 19.0 for Mac. First, to examine relations between all variables of interest, zero-order correlations were calculated. To investigate whether executive functions predicted report marks after correction for grade at the start of the study, sex, and LPE, separate multivariate GLM analyses (MANCOVA) were performed for each executive function score. Dependent variables were standardized report marks for Dutch, English, and mathematics. The following fixed factors and covariates were included: grade, sex, LPE (hereafter called demographic variables) and executive function score. After investigating main effects, we investigated interaction effects. To examine whether results were different for the different grades, we added the interaction term grade * executive function score. To investigate influence of sex, analyses were performed with inclusion of the interaction between sex and executive function score. Finally, we investigated a model with all executive function scores and all demographic variables to investigate their contribution together, and the same model without demographic variables to investigate the amount of variance predicted by executive function scores alone.

RESULTS

A total of 173 adolescents between 12.68 years and 18.05 years participated (age $M = 15.22$ years; $SD = 1.66$). Of those, 63.6% had highly educated parents, and the remainder had parents with a medium education level. **Table 1** shows outcomes on executive function measures and school performance, per grade and sex. Sex differences were seen on the Tower Test to the advantage of boys. On the other executive function measures, no sex differences were found. On all school report marks, there were differences between grades and between sexes: students from lower grades had higher report marks than students from higher grades, and girls achieved higher report marks than boys.

RELATION BETWEEN EXECUTIVE FUNCTIONS AND REPORT MARKS

Table 2 shows correlations between executive function measures and report marks. The BRI and MCI of the BRIEF-SR were the only executive function measures that significantly correlated with report marks. The BRI correlated with Dutch scores only ($r = -0.17$), while the MCI correlated with report marks in Dutch, English, and mathematics (between $r = -0.20$ to $r = -0.27$, $p < 0.05$). These correlations indicate that the more problems with behavior regulation a student reported, the lower the score for Dutch. In addition, the more problems with metacognition a student reported, the lower the score for Dutch, English, and mathematics.

SORTING TEST

MANCOVA analyses showed no significant main effect of Sorting Test score on report marks, $F_{(3,\ 165)} = 0.27$, $p = 0.847$, partial eta squared $= 0.01$. Repeating the analyses with interaction effects also showed no significant interaction effect between Sorting Test score and grade, and Sorting Test score and sex on report marks, resp. $F_{(6,\ 326)} = 1.07$, $p = 0.382$, partial eta squared $= 0.02$, and $F_{(3,\ 162)} = 0.58$, $p = 0.631$, partial eta squared $= 0.01$.

TOWER TEST

MANCOVA analyses showed no significant main effect of Tower Test score on report marks, $F_{(3,\ 165)} = 1.98$, $p = 0.119$, partial eta squared $= 0.04$. Repeating the analyses with the interaction effects also showed no significant interaction between Tower Test score and grade, and Tower Test score and sex on report marks, resp. $F_{(6,\ 326)} = 1.49$, $p = 0.181$, partial eta squared $= 0.03$ and $F_{(3,\ 162)} = 0.64$, $p = 0.588$, partial eta squared $= 0.01$.

BRIEF-SR BRI

MANCOVA analyses showed no significant main effect of BRI score on report marks, $F_{(3,\ 165)} = 1.99$, $p = 0.117$, partial eta squared $= 0.04$. Repeating the analyses with the interaction effects also showed no significant interaction between the score on the BRI and grade, and the BRI and sex on report marks, resp. $F_{(6,\ 326)} = 0.60$, $p = 0.729$, partial eta squared $= 0.01$, and $F_{(3,\ 162)} = 1.99$, $p = 0.118$, partial eta squared $= 0.04$.

BRIEF-SR MCI

MANCOVA analyses showed no significant main effect of MCI score on report marks, $F_{(3,\ 165)} = 2.12$, $p = 0.100$, partial eta squared $= 0.04$. Repeating the analyses with the interaction effects also showed no significant interaction between the score on the MCI and grade, and the MCI and sex on report marks, resp. $F_{(6,\ 326)} = 0.57$, $p = 0.751$, partial eta squared $= 0.01$, and $F_{(3,\ 162)} = 1.03$, $p = 0.382$, partial eta squared $= 0.02$.

MODEL WITH ALL EXECUTIVE FUNCTION SCORES

MANCOVA analyses with all executive function scores in one model showed no significant effects of any of the executive function scores [Sorting Test: $F_{(3,\ 162)} = 0.17$, $p = 0.918$, partial eta squared $= 0.00$; Tower Test: $F_{(3,\ 162)} = 2.36$, $p = 0.074$, partial eta squared $= 0.042$; BRIEF-SR BRI: $F_{(3,\ 162)} = 0.66$, $p = 0.577$, partial eta squared $= 0.01$; BRIEF-SR MCI: $F_{(3,\ 162)} = 1.34$, $p = 0.262$, partial eta squared $= 0.02$]. Investigating the three demographic variables showed that grade and sex were significant

Table 1 | Descriptive statistics of executive function measures and school performance, per grade and sex.

	Total	Grade 7		Grade 9		Grade 11		Grade effect	Sex effect
		Boys	Girls	Boys	Girls	Boys	Girls	Test statistic	Test statistic
N	173	32	28	30	33	20	30	–	–
Age	15.22 (1.66)	13.26 (0.43)	13.39 (0.35)	15.31 (0.36)	15.34 (0.30)	17.40 (0.37)	17.38 (0.40)	$F = 1615.31^{**}$	$F = 0.70$
LPE *(% high)*	63.6	62.5	57.1	73.3	60.6	55.0	70.0	–	–
Sorting test confirmed correct score[a]	10.75 (1.57)	10.47 (1.52)	10.32 (1.85)	10.73 (1.64)	11.03 (1.43)	10.55 (1.50)	11.30 (1.34)	$F = 2.17$	$F = 1.28$
Tower test total achievement score[a]	18.12 (3.43)	18.03 (3.13)	18.14 (3.16)	18.73 (3.49)	17.15 (2.79)	20.00 (4.81)	17.40 (3.10)	$F = 0.53$	$F = 6.03^{*}$
BRIEF–SR BRI[a]	54.99 (9.17)	56.59 (9.65)	54.50 (8.57)	51.17 (6.87)	55.88 (9.72)	58.00 (7.00)	54.57 (11.04)	$F = 1.10$	$F = 0.00$[b]
BRIEF–SR MCI[a]	64.79 (11.93)	64.47 (10.53)	61.43 (11.30)	63.13 (9.53)	63.72 (12.93)	74.00 (12.49)	64.93 (13.06)	$F = 4.12^{*}$	$F = 3.62$
Report mark Dutch	7.06 (0.83)	7.16 (0.91)	7.83 (0.70)	6.60 (0.53)	7.29 (0.66)	6.30 (0.57)	6.98 (0.75)		
Report mark Dutch standard score	0 (0.98)	0.04 (0.96)	0.99 (0.69)	−0.51 (0.66)	0.22 (0.87)	−0.92 (0.69)	−0.08 (0.92)	$F = 21.88^{**}$	$F = 45.16^{**}$
Report mark English	7.27 (1.15)	7.30 (0.98)	8.03 (1.11)	7.05 (0.96)	7.55 (0.95)	6.85 (0.82)	6.71 (1.47)		
Report mark English standard score	0 (0.98)	0.13 (0.82)	0.78 (0.84)	−0.18 (0.82)	0.19 (0.84)	−0.45 (0.68)	−0.60 (1.17)	$F = 17.60^{**}$	$F = 5.48^{*}$
Report mark mathematics	6.89 (1.11)	7.20 (1.14)	7.53 (0.83)	6.63 (0.92)	6.95 (0.95)	5.90 (0.91)	6.79 (1.25)		
Report mark mathematics standard score	0 (0.98)	0.26 (1.06)	0.60 (0.72)	−0.18 (0.76)	0.06 (0.84)	−0.91 (0.83)	−0.11 (1.09)	$F = 13.70^{**}$	$F = 9.79^{**}$

Values are M (SD). A dash indicates that the effect was not tested. Skewness of standard scores of report marks remained within an acceptable range (between −1 and +1). LPE, level of parental education; BRIEF-SR, Behavior Rating Inventory of Executive Function—Self-Report Version; BRI, Behavioral Regulation Index; MCI, Metacognition Index.

[a] *Raw score.*

[b] *Significant interaction effect between grade and sex, $F = 3.38^{*}$.*

$^{*}p < 0.05$, $^{**}p < 0.01$.

Table 2 | Zero-order correlations between background variables, executive function measures, and report marks.

	1	2	3	4	5	6	7	8	9	10
1. Group	1									
2. Sex	0.11	1								
3. LPE	0.04	−0.02	1							
4. Sorting Test confirmed correct score	0.16*	0.10	−0.01	1						
5. Tower Test total achievement score	0.04	−0.18*	−0.02	0.15	1					
6. BRIEF–SR BRI	0.01	0.00	−0.23**	−0.02	−0.01	1				
7. BRIEF–SR MCI	0.18*	−0.12	−0.08	0.06	0.15*	0.65**	1			
8. Standard score Dutch	−0.37**	0.39**	0.08	0.01	0.00	−0.17*	−0.27**	1		
9. Standard score English	−0.40**	0.12	−0.03	−0.06	0.01	−0.09	−0.20**	0.57**	1	
10. Standard score mathematics	−0.35**	0.18*	0.12	−0.02	0.10	−0.08	−0.20*	0.46**	0.50**	1

LPE, level of parental education; BRIEF-SR, Behavior Rating Inventory of Executive Function—Self-Report Version; BRI, Behavioral Regulation Index; MCI, Metacognition Index.

$^{*}p < 0.05$, $^{**}p < 0.01$.

[grade: $F_{(6,\ 326)} = 8.35$, $p < 0.001$, partial eta squared = 0.13; sex: $F_{(3,\ 162)} = 16.12$, $p < 0.001$, partial eta squared = 0.23], and LPE approached significance [$F_{(3,\ 162)} = 2.54$, $p = 0.058$, partial eta squared = 0.05]. In all analyses performed in this article, these three demographic variables were included and their effects were as described in this analysis. A model without demographic variables, but with all executive function scores, showed a significant effect for the BRIEF-SR MCI [$F_{(3,\ 166)} = 4.21$, $p = 0.007$, partial eta squared = 0.07]. This effect is smaller than the variance explained by demographic variables in the previous analysis.

DISCUSSION

The current study investigated whether executive functions predicted report marks in healthy adolescents aged 12–18 who were

secondary school students in the pre-university educational track. Results showed that performance on the Sorting Test and the Tower Test did not predict report marks for Dutch, English, and mathematics. There was a zero-order correlation between scores on the BRIEF-SR and report marks ($r = 0.17$–0.27). Such correlations are often reported in studies on the relation of executive function tests to school performance (e.g., St Clair-Thompson and Gathercole, 2006; Best et al., 2011). However, after correcting for grade, sex, and LPE, the BRIEF-SR did not predict report marks anymore. Moreover, sex did not influence the relation between executive functions and report marks, since the sex* executive function score interaction term was not significant in any of our models.

The magnitude of the correlation between the BRIEF-SR and report marks in the current study was comparable to that found in studies that controlled for intelligence (Latzman et al., 2010; Miller and Hinshaw, 2010). In the current study, all participants were in the pre-university educational track, the level at which the top 20% of Dutch students is studying (Ministry of Education, Culture and Science, 2009). By selecting this high-performing sample, we used a group that is relatively homogeneous with respect to potential and intellectual ability [estimated intelligence quotient (IQ) higher than 90; (Van den Bos et al., 2012)]. Students in the pre-university educational track are overrepresented in higher socioeconomic status (SES) groups and there is some evidence that these students are generally more mature with regard to neuropsychological development (Hackman and Farah, 2009). The advantage of our study is that it enabled some control over possible confounders related to SES, and over slow (neuro) psychological development due to lack of environmental support. The findings can be considered as strong for the upper segment of the educational system (pre-university education). On the other hand, our design has the drawback that the results cannot be generalized to all adolescents and that we were not able to take IQ-scores into account. Forthcoming studies will address executive functions in relation to boy-girl difference in students in the "vocational" educational track; given the fact that this group is characterized by broader variance in SES, intelligence, learning motivation, and study results, findings are anticipated to be different than reported in the present paper. This would be relevant in terms of applicability of the findings in educational practice, notably in councelling of students and their parents and teachers with regard of the development of executive functions.

The fact that the BRIEF-SR did relate to report marks, while the objective neuropsychological tests (Sorting Test and Tower Test) did not, may illustrate the higher ecological validity of the BRIEF-SR. Yet, controlling for grade, sex, and LPE removed the relation between the BRIEF-SR and report marks. Thus, if one knows whether a student in a preuniversity educational track is a boy or a girl, which grade the student is in, and the educational level of the parents, one can predict report marks as well as with the score on the BRIEF-SR. This indicates that the BRIEF-SR measured aspects of performance in school that could be explained by grade, sex, and LPE.

Why could the executive function tests in the present study not predict school performance any better than grade, sex, and LPE? One explanation has to do with the fact that our study used a sample which was homogeneous with respect to educational track: only students from pre-university level were investigated, i.e., 20% of all students in secondary education. This choice will have reduced the influence of the background variables on the relation between executive functions and school performance. Our findings are in line with those of Willoughby et al. (2012), a study that resembles the present study by taking into account more than only measured confounders. Moreover, other studies that did not find any effects between executive functions and school performance may have remained unpublished. Another explanation is that students who are elected for the preuniversity track in secondary education are more mature in executive functioning in comparison to students in other educational tracks. This is of importance since—usually—, primary schools advise each student which educational level in secondary education is appropriate for them. This advice takes into account not only cognitive performance, but also expected development, motivation for school, and study approach (Driessen, 2005). It could be that, unknowingly, executive functions are taken into account as well. Students with good executive skills would then be advised to go to preuniversity education, while students with poorer executive skills would be advised to go to general secondary education or prevocational education. Future research should investigate the relation between executive functions and school performance in general secondary education and prevocational education as well. Possibly, effects will be found at these other educational levels.

However, also within preuniversity education, students themselves report that they differ with respect to their executive function skills (Coenen et al., 2011). Moreover, sex differences in self-control, which is closely linked to executive functions, appear to contribute to sex differences in school performance (Downey et al., 2005; Duckworth and Seligman, 2006; Hyde et al., 2007; Steinmayr and Spinath, 2008). This may indicate that the executive function tests used in this study were not sensitive enough to measure differences in high-performing healthy adolescents. Executive function tests used in clinical practice are often not sensitive enough to distinguish executive function difficulties in clinical groups (Chan et al., 2008), let alone in healthy subjects. Since each executive function test also measures other (not executive) functions, so called task-impurity, this may trouble the accurate measurement of executive functions (Miyake and Friedman, 2012). To accurately measure differences in executive functions between healthy high-performing adolescents, other tests, a combination of tests, or statistical methods such as the latent variable approach may be needed (Miyake and Friedman, 2012).

A strong point of the current study is that it used report marks to estimate school performance. Most studies measure school performance with standardized tests. An advantage of standardized tests is that these tests are similar for all participants in the study (OECD, 2007). A disadvantage of standardized tests is the lack of ecological validity, because standardized tests are not the outcomes on which students are being assessed in school (Cohen et al., 1995; Wolfson and Carskadon, 2003). Students are reliant on report marks for their school success, as report marks indicate whether a student may pass to the next grade or enter a certain school or educational track. Report marks may also have higher reliability than standardized tests because they involve

multiple measurements and more closely measure learning that takes place at school (Wolfson and Carskadon, 2003). Thus, report marks may give a better estimation of real life outcomes than standardized tests.

The present study shows that school performance in healthy, high-performing adolescents could not be predicted by scores on the Sorting Test, Tower Test, and BRIEF-SR. It raises the question whether and to what extent school performance in this sample depends on executive functions. In this sample of healthy, high-performing adolescents, school performance may be affected more strongly by other cognitive factors, for example, content knowledge of the school subjects, or psychological factors such as motivation or personality. Moreover, the study illustrates that controlling for confounders is very important in research on the effect of executive functions on school performance (Willoughby et al., 2012). Future research may investigate whether these results also hold for other executive function tests and other samples. For instance, are similar results seen in adolescents who study at other educational levels? And in adolescents who repeated a grade or have a developmental disorder? Based on the current study, we can conclude that the executive functions measured with the Sorting Test, Tower Test, and BRIEF-SR do not play a major role in report marks obtained by healthy, high-performing adolescents.

AUTHOR CONTRIBUTIONS

All authors contributed to the design and execution of the study. All authors drafted and revised the study, and approved the final version. All authors agreed to be accountable for all aspects of the work.

REFERENCES

Alvarez, J. A., and Emory, E. (2006). Executive function and the frontal lobes: a meta-analytic review. *Neuropsychol. Rev.* 16, 17–42. doi: 10.1007/s11065-006-9002-x

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., and Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learn. Individ. Differ.* 21, 327–336. doi: 10.1016/j.lindif.2011.01.007

Chan, R. C. K., Shum, D., Toulopoulou, T., and Chen, E. Y. H. (2008). Assessment of executive functions: review of instruments and identification of critical issues. *Arch. Clin. Neuropsychol.* 23, 201–216. doi: 10.1016/j.acn.2007.08.010

Checa, P., Rodriguez-Bailon, R., and Rueda, M. (2008). Neurocognitive and temperamental systems of self-regulation and early adolescents' social and academic outcomes. *Mind Brain Educ.* 2, 177–187. doi: 10.1111/j.1751-228X.2008.00052.x

Checa, P., and Rueda, M. R. (2011). Behavioral and brain measures of executive attention and school competence in late childhood. *Dev. Neuropsychol.* 36, 1018–1032. doi: 10.1080/87565641.2011.591857

Clark, M. A., Thompson, P., and Vialle, W. (2008). Examining the gender gap in educational outcomes in public education: involving pre-service school counsellors and teachers in cross-cultural and interdisciplinary research. *Int. J. Adv. Couns.* 30, 52–66. doi: 10.1007/s10447-007-9044-7

Coenen, J., Meng, C., and Van Der Velden, R. (2011). *Schoolsucces van Jongens en Meisjes in het HAVO en VWO: Waarom Meisjes Het Beter Doen (No. ROA-R-2011/2). nieuws.wegenerweb.nl*. Maastricht: Researchcentrum voor Onderwijs en Arbeidsmarkt (ROA).

Cohen, G., Bronson, M., and Casey, M. (1995). Planning as a factor in school achievement. *J. Appl. Dev. Psychol.* 16, 405–428. doi: 10.1016/0193-3973(95)90027-6

Dekker, S., Krabbendam, L., Aben, A., De Groot, R., and Jolles, J. (2013a). Coding task performance in early adolescence: a large-scale controlled study into boy-girl differences. *Front. Psychol.* 4:550. doi: 10.3389/fpsyg.2013.00550

Dekker, S., Krabbendam, L., Lee, N. C., Boschloo, A. M., De Groot, R., and Jolles, J. (2013b). Sex differences in goal orientation in adolescents aged 10-19: the older boys adopt work-avoidant goals twice as often as girls. *Learn. Individ. Differ.* 26, 196–200. doi: 10.1016/j.lindif.2012.07.011

Delis, D. C., Kaplan, E., and Kramer, J. H. (2001). *Delis-Kaplan Executive Function System, Technical Manual*. San Antonio, TX: Psychological Corporation.

Downey, D., Yuan, V., and Anastasia, S. (2005). Sex differences in school performance during high school: puzzling patterns and possible explanations. *Soc. Q.* 46, 299–321. doi: 10.1111/j.1533-8525.2005.00014.x

Driessen, G. (2005). De totstandkoming van de adviezen voortgezet onderwijs: invloeden van thuis en school. [The coming about of the recommendation for secondary school]. *Pedagogiek* 25, 279–298.

Driessen, G., and Van Langen, A., (2010). *De Onderwijsachterstand van Jongens: Omvang, Oorzaken en Interventies [The Educational Disadvantage of Boys: Magnitude, Causes and Interventions]*. Nijmegen: ITS: Radboud Universiteit Nijmegen.

Duckworth, A., and Seligman, M. (2006). Self-discipline gives girls the edge: gender in self-discipline, grades, and achievement test scores. *J. Educ. Psychol.* 98, 198–208. doi: 10.1037/0022-0663.98.1.198

Giedd, J. N. (2008). The teen brain: insights from neuroimaging. *J. Adolesc. Health* 42, 335–343. doi: 10.1016/j.jadohealth.2008.01.007

Gioia, G. A., and Isquith, P. K. (2004). Ecological assessment of executive function in traumatic brain injury. *Dev. Neuropsychol.* 25, 135–158. doi: 10.1080/87565641.2004.9651925

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D. K., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179. doi: 10.1073/pnas.0402680101

Guy, S. C., Isquith, P. K., and Gioia, G. A. (2004). *BRIEF-SR: Behavior Rating Inventory of Executive Function-Self-Report Version: Professional Manual*. Lutz, FL: Psychological Assessment Resources, Inc.

Hackman, D. A., and Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends. Cogn. Sci.* 13, 65–73. doi: 10.1016/j.tics.2008.11.003

Hofmann, W., Schmeichel, B. J., and Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends Cogn. Sci.* 16, 174–180. doi: 10.1016/j.tics.2012.01.006

Huizinga, M., Dolan, C. V., and Van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Hyde, J., Lindberg, S., and Wilson, A. (2007). "Facts and assumptions about the nature of gender differences and the implications for gender equity," in *Handbook for Achieving Gender Equity Through Education*, eds S. S. Klein, B. Richardson, D. A. Grayson, L. H. Fox, C. Kramarae, D. S. Pollard, and C. A. Dwyer (Mahwah, NJ: Lawrence Erlbaum Associates, Inc), 19–32.

Jimerson, S. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *Sch. Psychol. Rev.* 30, 420–437.

Latzman, R. D., Elkovitch, N., Young, J., and Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *J. Clin. Exp. Neuropsycol.* 32, 455–462. doi: 10.1080/13803390903164363

Lenroot, R. K., and Giedd, J. N. (2010). Sex differences in the adolescent brain. *Brain Cogn.* 72, 46–55. doi: 10.1016/j.bandc.2009.10.008

Luciana, M., Collins, P. F., Olson, E. A., and Schissel, A. M. (2009). Tower of London performance in healthy adolescents: the development of planning skills and associations with self-reported inattention and impulsivity. *Dev. Neuropsychol.* 34, 461–475. doi: 10.1080/87565640902964540

Machin, S. (2005). Gender and student achievement in english schools. *Oxford Rev. Econ. Pol.* 21, 357–372. doi: 10.1093/oxrep/gri021

Martens, R., Hurks, P. P. M., Meijs, C., Wassenberg, R., and Jolles, J. (2011). Sex differences in arithmetical performance scores: central tendency and variability. *Learn. Individ. Differ.* 21, 549–554. doi: 10.1016/j.lindif.2011.06.003

Miller, M., and Hinshaw, S. P. (2010). Does childhood executive function predict adolescent functional outcomes in girls with ADHD? *J. Abnorm. Child Psychol.* 38, 315–326. doi: 10.1007/s10802-009-9369-2

Miller, M., Nevado-Montenegro, A. J., and Hinshaw, S. P. (2012). Childhood executive function continues to predict outcomes in young adult females with and without childhood-diagnosed ADHD. *J. Abnorm. Child Psychol.* 40, 657–658. doi: 10.1007/s10802-011-9599-y

Ministry of Education, Culture and Science. (2006). *Besluit Vaststelling Kerndoelen Onderbouw VO en Aanpassing INRICHTINGSBESLUIT WVO [Decision Establishing core Objectives for the Lower School in Secondary Education and Amending the Decision Implementing the Law on Secondary Education] (Kamerstuk WJZ/2006/24505(3805))*. Available online at: http://www.rijksoverheid.nl/bestanden/documenten-en-publicaties/kamerstukken/2006/06/16/besluit-vaststellingkerndoelen-onderbouw-vo-en-aanpassing-inrichtingsbesluit-wvo/24506a.pdf (Accessed: May/20/2010).

Ministry of Education, Culture and Science. (2009). *Kerncijfers 2004-2008: Onderwijs, Cultuur en Wetenschap* [*Core figures 2004-2008: Education, Culture and Science*]. (OCW39.008/1.000/08BK2009B003). Den Haag: Ministry of Education, Culture and Science.

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol.* 21, 8–14. doi: 10.1177/0963721411429458

OECD. (2007). *PISA 2006. Vol. 1: Analysis.* Paris: OECD. doi: 10.1787/9789264040014-en

Olson, K., Jacobson, K. K., and Van Oot, P. (2013). Ecological validity of pediatric neuropsychological measures: current state and future directions. *Appl. Neuropsychol. Child* 2, 17–23. doi: 10.1080/21622965.2012.686330

Reed, H., Ouwehand, C., Van der Elst, W., Boschloo, A., and Jolles, J. (2010). Measures of school performance for use in educational neuropsychology. *Front. Neurosci. Conference Abstract EARLI SIG22 - Neuroscience and Education*. doi: 10.3389/conf.fnins.2010.11.00058

St Clair-Thompson, H. L., and Gathercole, S. E. (2006). Executive functions and achievements in school: shifting, updating, inhibition, and working memory. *Q. J. Exp. Psychol.* 59, 745–759. doi: 10.1080/17470210500162854

Steenbergen-Hu, S., and Moon, S. M. (2011). The effects of acceleration on high-ability learners: a meta-analysis. *Gifted Child Q.* 55, 39–53. doi: 10.1177/0016986210383155

Steinmayr, R., and Spinath, B. (2008). Sex differences in school achievement: what are the roles of personality and achievement motivation? *Eur. J. Personality* 22, 185–209. doi: 10.1002/per.676

Van den Bos, W., Crone, E. A., and Güroğlu, B. (2012). Brain function during probabilistic learning in relation to IQ and level of education. *Dev. Cogn. Neurosci.* 2s, s78–s89. doi: 10.1016/j.dcn.2011.09.007

Van Langen, A., Bosker, R., and Dekkers, H. (2006). Exploring cross-national differences in gender gaps in education. *Educ. Res. Eval.* 12, 155–177. doi: 10.1080/13803610600587016

Veenman, M. V. J., Kok, R., and Blöte, A. W. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instr. Sci.* 33, 193–211. doi: 10.1007/s11251-004-2274-8

Willoughby, M. T., Kupersmidt, J. B., and Voegler-Lee, M. E. (2012). Is preschool executive function causally related to academic achievement? *Child Neuropsychol.* 18, 79–91. doi: 10.1080/09297049.2011.578572

Wolfson, A. R., and Carskadon, M. A. (2003). Understanding adolescents' sleep patterns and school performance: a critical appraisal. *Sleep Med. Rev.* 7, 491–506. doi: 10.1016/S1087-0792(03)90003-7

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 27 October 2013; accepted: 20 March 2014; published online: 08 April 2014. Citation: Boschloo A, Krabbendam L, Aben A, de Groot R and Jolles J (2014) Sorting Test, Tower Test, and BRIEF-SR do not predict school performance of healthy adolescents in preuniversity education. Front. Psychol. **5**:287. doi: 10.3389/fpsyg.2014.00287 This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.*

Copyright © 2014 Boschloo, Krabbendam, Aben, de Groot and Jolles. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 07 May 2014
doi: 10.3389/fpsyg.2014.00390

Executive control training from middle childhood to adolescence

***Julia Karbach*[1]* and *Kerstin Unger*[2]**

[1] Department of Educational Science, Saarland University, Saarbrücken, Germany
[2] Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, USA

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
M. Rosario Rueda, Universidad de Granada, Spain
Katharina Zinke, University of Tübingen, Germany
Kristina Küper, Leibniz Research Centre for Working Environment and Human Factors, Germany

***Correspondence:**
Julia Karbach, Department of Educational Science, Saarland University, Campus A4 2, D-66123 Saarbrücken, Germany
e-mail: j.karbach@mx.uni-saarland.de

Executive functions (EFs) include a number of higher-level cognitive control abilities, such as cognitive flexibility, inhibition, and working memory, which are instrumental in supporting action control and the flexible adaptation changing environments. These control functions are supported by the prefrontal cortex and therefore develop rapidly across childhood and mature well into late adolescence. Given that executive control is a strong predictor for various life outcomes, such as academic achievement, socioeconomic status, and physical health, numerous training interventions have been designed to improve executive functioning across the lifespan, many of them targeting children and adolescents. Despite the increasing popularity of these trainings, their results are neither robust nor consistent, and the transferability of training-induced performance improvements to untrained tasks seems to be limited. In this review, we provide a selective overview of the developmental literature on process-based cognitive interventions by discussing (1) the concept and the development of EFs and their neural underpinnings, (2) the effects of different types of executive control training in normally developing children and adolescents, (3) individual differences in training-related performance gains as well as (4) the potential of cognitive training interventions for the application in clinical and educational contexts. Based on recent findings, we consider how transfer of process-based executive control trainings may be supported and how interventions may be tailored to the needs of specific age groups or populations.

Keywords: executive control, cognitive training, childhood, adolescence, cognitive plasticity

INTRODUCTION

Over the last decade, the scientific interest in cognitive interventions designed to improve cognitive functions in childhood and adolescence has been rapidly increasing. The many studies investigating the benefits of cognitive training interventions showed that cognitive plasticity is considerable not only in children and adolescents, but also up to old age (for recent reviews, see Buitenweg et al., 2012; Diamond, 2012; Karbach and Schubert, 2013; Kray and Ferdinand, 2013; Strobach et al., 2014; Titz and Karbach, 2014; Verhaeghen, 2014). These studies usually showed significant performance improvements on the trained tasks. Moreover, they oftentimes also revealed near transfer to tasks that were not explicitly trained but measured the same construct as the training task, and sometimes even far transfer to tasks measuring a different construct.

Despite these encouraging findings, the literature clearly shows that these transfer effects were not consistent across studies, a fact that has inspired intense recent debates regarding the transferability of training-induced performance gains (e.g., Shipstead et al., 2012; Melby-Lervåg and Hulme, 2013; Redick et al., 2013). The inconsistent pattern of results may be explained by the large differences in terms of the type, intensity, and duration of the training regimes and the fact that different methodologies haven been adopted across studies. Thus, the comparability of previous results is often very limited.

In addition, it makes sense to differentiate different types of cognitive training interventions: strategy-based training refers to interventions involving the training of task-specific approaches designed to support the execution of certain tasks. It has often been applied in memory training studies and typical examples include mnemonic techniques, such as the method of loci. This type of memory strategy training often resulted in large and often long-lasting improvements on the training task, but induced only limited transfer (for meta-analyses, see Verhaeghen et al., 1992; Rebok et al., 2007). Multi-domain training interventions are usually more complex and engage multiple cognitive processes (e.g., game-based training), yielding broad but often small transfer effects (e.g., Basak et al., 2008). The main disadvantage of multi-domain trainings is that their complex nature makes it hard to determine which specific features of the training regime induced transfer.

In contrast, process-based training protocols are not task-specific because they target more general processing capacities supporting a range of cognitive operations, such as speed of processing or executive functions (EFs). Some process-based interventions, mainly from the domain of EF, have resulted in very promising widespread transfer across the lifespan (Hertzog et al., 2008; Karbach and Schubert, 2013; Kray and Ferdinand, 2013; Titz and Karbach, 2014), suggesting that process-based training might be more efficient than strategy-based interventions. The fact that

EF may be improved by means of cognitive training is of particular importance in childhood and adolescence, because EF is a strong predictor for various life outcomes, such as academic attainment, socioeconomic status, and physical health (e.g., Eigsti et al., 2006; Blair and Razza, 2007; Moffitt et al., 2011). Moreover, behavioral and neural plasticity is particularly high in childhood and the brain areas serving EF (i.e., the prefrontal lobes) are especially sensitive to environmental influences in children (cf. Bull et al., 2011). It is therefore not surprising that numerous training interventions have been designed to improve executive functioning across the lifespan, many of them targeting children and adolescents. These studies have included normally developing children as well as individuals suffering from neurodevelopmental or psychiatric disorders, some of which are characterized by significant cognitive deficits [e.g., attention-deficit hyperactivity disorder (ADHD) or autism].

A number of recent systematic reviews and meta-analyses have focused on training interventions targeting EF in children. Some of them have analyzed findings from samples with cognitive impairments (e.g., Rapport et al., 2013; Chacko et al., 2014), others have selectively focused on specific types of training (e.g., Kray and Ferdinand, 2013) or on specific methodological approaches, such as neuroscientific techniques (e.g., Buschkuehl et al., 2012; Jolles and Crone, 2012), or on specific outcome measures, such as academic achievement (e.g., Titz and Karbach, 2014). Comprehensive reviews including training on different components of EF in samples of normally developing children have been focused on preschoolers (e.g., Diamond, 2012; Zelazo and Lyons, 2012) and so far a systematic review of recent findings on EF training in middle childhood and adolescence is still missing. Such a review may contribute to the understanding of the cognitive mechanisms underlying plasticity of cognitive functions across their development in middle childhood and adolescence. Considering the importance of EF for numerous life outcomes, the identification of successful cognitive training interventions may not only be beneficial for the compensation of cognitive deficits in clinical samples, but also to promote cognitive performance and development in healthy children and adolescents.

Therefore, the aim of this review is to (1) illustrate the concept of EF, its neural correlates and age-related changes in middle childhood and adolescence as an introduction to (2) the presentation of selected recent findings of process-based EF training in this age group, followed by (3) the description of individual differences in training-related improvements. We close by (4) outlining potential applications of EF training in clinical and educational settings.

DEFINITIONS OF EF

The term executive control refers to a broad collection of higher-order cognitive functions that allow individuals to flexibly regulate their thoughts and actions in the service of adaptive, goal-directed behavior. EFs are typically thought to encompass a wide range of mental processes that vary in complexity and abstractness, such as working memory, cognitive flexibility, attentional control, planning, concept formation, or feedback processing (e.g., Jurado and Rosselli, 2007). Working memory serves to update and monitor information and to code task-relevant information. This relevant information is held in working memory until it is no longer needed and subsequently replaced with newer, more relevant information. Working memory is required to mentally relate, integrate, and recombine information across different time scales and hence plays a pivotal role for more complex EFs such as planning or concept formation. Following a conversation in a foreign language puts high demands on working memory resources, as does difficult mental arithmetic or planning the optimal route from city center to airport during rush hour traffic. Also referred to as *shifting*, *attention switching*, or *task switching*, cognitive flexibility refers to the ability to flexibly shift between tasks, goals, or mental sets. It involves disengaging from currently irrelevant information (i.e., the previous task set) and focusing on currently relevant information (i.e., the upcoming task; e.g., Meiran, 1996; Monsell, 2003). Cognitive flexibility allows us to think divergently and creatively and to respond quickly to unpredicted changes in the environment. It helps us to change the perspective and develop new solution ideas when we are stuck with a problem (e.g., trying to handle a new electronic device or software tool) or to use unexpected opportunities such as backing up in a parking spot that suddenly opens up behind us while we are waiting for another car to leave a parking space in front of us. Attentional control is required when we need to focus on a specific stimulus while minimizing interference from irrelevant stimuli. In everyday life, we use this ability when we are talking on the phone and have to tune out conversations of other people around us. Another form of control involves the inhibition of automatic or impulsive response tendencies and unwanted emotions. For example, if a deer suddenly jumps out in front of our car, we have to suppress the tendency to swerve. Similarly, if we want to lose weight, we have to resist sweets and fatty foods and social norms dictate us not to yell at another person even if we are angry.

The heterogeneity of the processes described above highlights the need to determine the structure and organization of executive control more precisely. A key question in this context was whether EFs are best characterized as unitary or multi-dimensional in nature. Early theoretical frameworks mostly adopted the perspective that a common cognitive mechanism or ability underlies executive functioning. Prominent examples are Norman and Shallice's (1986) "Supervisory Attentional System" or the closely related "Central Executive" in Baddeley's (1986) working memory model. A more recent proposal by Duncan et al. (1996) established a theoretical link between the concept of a prefrontally based unitary control system and Spearman's general intelligence factor *g* (see also Denckla and Reiss, 1997; Kimberg et al., 1997; de Frias et al., 2006). In a similar vein, Salthouse (2005) noted that interindividual differences in executive functioning may tap basic reasoning skills and perceptual speed (but see Ardila et al., 2000; Friedman et al., 2006). Empirical support for the unitary nature of executive control comes from psychometric studies showing that different components are substantially correlated at the latent variable level (e.g., Miyake et al., 2000; Friedman et al., 2008, 2011). The intercorrelations among these latent factors, however, are usually moderately high, indicating that EFs comprise clearly separable subcomponents even though they may share some commonalities.

In an influential study, Miyake et al. (2000) provided first conclusive evidence for the "unity/diversity" framework by applying a latent variable approach to a task battery designed to capture three putative core components of executive control (see above for a more detailed description of the involved processes): (1) flexibly switching between different task sets or mental representations (*shifting*), (2) updating, removing, and monitoring working memory contents (*updating*), and (3) overriding prepotent response tendencies and suppressing attention to irrelevant stimuli as well as unwanted thoughts and emotions (*inhibition*; Miyake et al., 2000). The authors demonstrated that a full three-factor model that allowed correlations between the three latent variables yielded a better fit than either a three-factor model that did not allow for such correlations or any other single- or two-factor model. Interestingly, subsequent work showed that when variance common to all executive tasks was accounted for by a unity factor (referred to as common EF), only the shifting and updating factors captured unique variance (Friedman et al., 2011). Thus, there was no evidence for an inhibition-specific ability that is separable from the common factor. As Miyake and Friedman (2012) pointed out, a strong candidate mechanism for this common basic ability is the stable maintenance of task goals and goal-relevant representations in working memory, whereas the updating- and shifting-specific component might reflect effective gating and clearance of those representations (Herd et al., submitted). Specifically, it has been hypothesized that updating might be associated with efficient gating of information into working memory and/or controlled long-term memory retrieval (Miyake and Friedman, 2012). The shifting-specific component, in contrast, has been suggested to reflect mental "stickiness" (Altamirano et al., 2010; Herd et al., submitted), denoting the uncontrolled, automatic persistence of goal representations that are no longer relevant and hence should be removed from working memory.

NEURAL UNDERPINNINGS OF EF

Historically, the study of the neural substrates underpinning EFs originated from the observation of common deficits in patients with frontal lobe lesions (Stuss and Benson, 1986), including impairments in working memory, planning, and inhibition (Shallice and Burgess, 1991). Although the prefrontal cortex (PFC) is thought to play a key role in mediating executive control, neuroimaging and lesion studies demonstrated that the performance of executive tasks is associated with activation in a large set of brain regions, involving prefrontal and parietal areas, motor regions, as well as subcortical structures, such as basal ganglia and thalamus (Duncan and Owen, 2000; Dosenbach et al., 2008; Niendam et al., 2012).

In line with the unity/diversity framework, a number of reviews and meta-analyses demonstrated that performance on different EF tasks reflects the joint contribution of a common frontoparietal network and unique, component-specific brain regions (Wager and Smith, 2003; Wager et al., 2004; Collette et al., 2006; Niendam et al., 2012). Specifically, it has been shown that shifting, updating, and inhibition tasks elicit overlapping activation in frontal [e.g., dorsolateral PFC (DLPFC), the anterior cingulate cortex (ACC)] and parietal regions (e.g., superior and inferior parietal lobe, precuneus) associated with the common executive control network. Component-specific (i.e., non-overlapping) activations were observed in distinct prefrontal, occipital and temporal areas (including BAs 6, 10, 11, 19, 13, and 37). Furthermore, analyses showed unique activation patterns in subcortical regions, including caudate, thalamus, putamen, and cerebellum, for inhibition and updating tasks.

Similar conclusions have been drawn in a positron emission tomography (PET) study by Collette et al. (2005) that used conjunction analyses to identify common neural substrates of the executive tasks administered by Miyake et al. (2000). Findings revealed that left superior parietal gyrus, right intraparietal gyrus, right intraparietal sulcus and, albeit less robustly, left middle and inferior frontal gyri were commonly engaged by all three executive processes. Although pairwise comparisons of the specific component processes showed dissociations in frontoparietal activation patterns, the observed differences do not easily map onto the latent factor structure suggested by Miyake et al. (2000) and Miyake and Friedman (2012).

Consistent with Miyake et al.'s (2000) proposal, the common frontoparietal network, especially the prefrontal part, is thought to play a major role in actively representing and maintaining task-goals, task context or task sets (rules) in order to bias downstream information processing (Miller and Cohen, 2001; Rossi et al., 2009). Munakata et al. (2011) recently proposed a framework that describes how inhibitory control emerges from this key function of the PFC. Specifically, the authors argue against the widely held view that certain prefrontal regions, such as the right inferior frontal gyrus, are functionally specialized to subserve inhibition. Instead, the framework posits that specific contributions of different prefrontal regions to inhibitory processes depend on the kind of information they represent and their interconnections with other brain areas. Thus, for instance, the ACC and related medial frontal areas are thought to use signals of conflict, errors, or uncertainty to inhibit inappropriate motor responses via projections to the subthalamic nucleus.

A prevalent view is that parietal regions such as intraparietal sulcus or inferior parietal lobule are involved in the top-down control of attention (Corbetta and Shulman, 2002) and may support executive control by subserving functions such as cue decoding or signaling of stimulus conflict (Dosenbach et al., 2008). In addition, parietal activation has been linked to maintenance of stimulus–response (S–R) mappings (Bunge, 2004) well as manipulation of working memory contents (Wendelken et al., 2008).

Furthermore, accumulating evidence indicates that EFs critically rely on complex interactions between PFC and subcortical structures via frontal corticobasal ganglia and frontal corticocerebellar circuits (e.g., Heyder et al., 2004; Gruber et al., 2006; O'Reilly and Frank, 2006). For instance, Miyake and Friedman (2012) suggested that updating might be associated with selective and efficient gating of information into working memory via corticostriatal loops. In addition, Herd et al. (submitted) used a computational modeling approach to explore two potential neural mechanisms underlying individual differences in the shifting-specific component of executive control: (1) recurrent connection strength in PFC and (2) efficient clearing of old representations from working memory upon gating decisions from

the basal ganglia. Both are thought to influence the tendency to maintain information in working memory that is no longer task-relevant.

AGE-RELATED CHANGES IN EF

Infant research has shown that elementary forms of executive control emerge within the first year of life (Carpenter et al., 1998; Diamond, 2006). Although core components of executive control, including working memory, inhibition and attentional flexibility, can be observed in preschoolers as young as 3 years of age (Hughes, 1998), EFs continue to improve throughout childhood into late adolescence or even adulthood (Davidson et al., 2006; Huizinga et al., 2006; Diamond, 2013). In recent years, a number of studies have addressed the key question of whether the unity/diversity framework appropriately describes the structure of EFs in children and adolescents (e.g., Lehto et al., 2003; Gathercole et al., 2004; Huizinga et al., 2006; Wiebe et al., 2008, 2011; Rose et al., 2011). Most findings support the notion that the latent factor structure of executive control changes qualitatively across development, from a unitary structure (i.e., a single-factor structure) in preschoolers to multiple subcomponents in school-age children and adolescents.

Developmental trajectories of EFs are thought to be inextricably linked to maturational changes of prefrontal regions and associated cortical and subcortical structures, including parietal regions and basal ganglia (e.g., Casey et al., 2005; Bunge and Wright, 2007; Luna et al., 2010). Behavioral improvements in cognitive control coincide with synaptic pruning and increased myelination as well as experience-dependent synaptic strengthening (Sowell et al., 2001; Bjorklund, 2005; Dawson and Guare, 2010). Some regions within PFC, such as orbitofrontal cortex, reach structural maturity at an earlier age, whereas others, such as DLPFC, show more protracted maturational time course (Gogtay et al., 2004). There is evidence to suggest that those differences in structural maturation are paralleled by changes in functional maturation and hence may account for distinct developmental trajectories among EFs (Bunge and Zelazo, 2006). Moreover, it has been demonstrated that there are substantial developmental changes in the structure of neural network(s) underlying executive control (Fair et al., 2008), with the number of short-range connections decreasing (segregation) and the number long-range connections increasing (integration) from childhood to adulthood. Using neural network modeling, Edin et al. (2007) showed that age-related differences in activation in the superior frontal sulcus and intraparietal sulcus during working tasks can be accounted for by stronger fronto-parietal (i.e., interregional). In contrast, stronger intraregional connectivity as well as faster conduction or increased coding specificity could not explain developmental changes in patterns of brain activity. In the following, we outline major developmental changes of the three core EFs cognitive flexibility/shifting, working memory/updating, and inhibition.

The ability to flexibly shift between task sets shows the most protracted development and continues to improve into adolescence (Chevalier and Blaye, 2009; Best and Miller, 2010; Diamond, 2013). Although 3- and 4-year-old children are able to successfully shift between two simple rules (e.g., Zelazo, 2004; Moriguchi and Hiraki, 2009, 2011), performance continues to improve at later ages for more complex task sets and higher numbers of rules. Several studies have consistently shown that two components of task shifting – the ability to switch from one rule to another rule (i.e., switching *per se*) and the ability to maintain and select two (or more) rules – follow different developmental time courses (e.g., Crone et al., 2004, 2006; Kray et al., 2004, 2008, 2012; Huizinga and van der Molen, 2007; Karbach and Kray, 2007). For instance, Huizinga and van der Molen (2007) reported that children's set switching abilities reached adult levels by the age of 11 years, whereas set maintenance continued to improve by the age of 15 years. Moreover, findings by Crone et al. (2006) indicated that the different developmental trajectories of rule representation/retrieval and rule switching/suppression are associated with differences in the recruitment of ventrolateral PFC (VLPFC) and pre-supplementary motor area (pre-SMA/SMA), respectively. Another study on task switching (Rubia et al., 2006) found age-related increases in the recruitment of several brain regions that have been implicated in shifting, including right inferior PFC, left parietal cortex, ACC, and striatum. In contrast, Wendelken et al. (2012) observed similar activation of frontoparietal control regions across children and adults during task switching.

While basic updating processes can be observed in 9- to 12-month-old infants, the ability to manipulate items in working memory develops later and over a longer time range (Diamond, 2013). Working memory performance at more complex tasks has been shown to improve linearly from pre-school age to adolescence (Gathercole et al., 2004), with age differences varying as a function of complexity (Luciana et al., 2005). Developmental neuroimaging studies typically focused on simple maintenance demands (Bunge and Wright, 2007). These studies revealed a complex pattern of age-related changes in brain activation. Some of the regions that have been associated with working memory processes in adults, such as the superior frontal sulcus and the intraparietal sulcus, show activation increases across childhood and adolescence, whereas others, such as the DLPFC and parietal cortices, are recruited to a lesser degree. There is also evidence for qualitative changes in neural activation. Scherf et al. (2006) observed that children engaged a compensatory network, including caudate nucleus, anterior insula, and lateral cerebellum, whereas adolescents recruited an adult-like working-memory circuitry comprising core structures like DLPFC and ACC, albeit to a lesser degree. Findings from other studies indicated that children are less able to suppress interference (Kray et al., 2012). Furthermore, Crone et al. (2006) provided evidence that children's performance deficits at tasks that require manipulation of information in working might be related to their failure to recruit frontoparietal regions.

Inhibitory control develops rapidly during the preschool years and typically continues to improve into middle childhood (Kray et al., 2009, 2012; Best and Miller, 2010). However, some studies using computerized tasks reported continued improvement until adolescence or even young adulthood (e.g., Huizinga et al., 2006). Developmental trajectories have been found to depend strongly of the nature of the inhibition task, suggesting that different tasks tap into distinct control processes (Nigg, 2000). Similar to the developmental course of shifting and working memory,

age differences in inhibition vary as a function of task (rule) complexity (Zelazo, 2006). Depending on the specific inhibitory requirements of the task, the ability to override a prepotent response has been found to improve most rapidly between ages 3 and 4 years (Hughes, 1998) or – particularly when the task involves concurrent working memory demands, the response bias is stronger, or responses have to be inhibited at a late stage (of execution) – between ages 5 and 8 years (Romine and Reynolds, 2005). Best and Miller (2010) suggested that early improvements in inhibitory control mainly reflect qualitative changes in information processing such as children's conceptual understanding of the hierarchical rule system underlying tasks like the dimensional change card sort (DCCS; Zelazo, 2006), while later improvements indicate quantitative changes, such as increasing efficiency of the underlying cognitive mechanism. Neuroimaging studies established a functional link between recruitment of right VLPFC as well as functionally connected subcortical areas such as thalamus, nucleus caudate, cerebellum, and the development of mature response control (Rubia et al., 2001; Bunge et al., 2002). Moreover, evidence from EEG studies indicated that refinements in stimulus processing (e.g., better stimulus discrimination) contribute to age-related performance increments on inhibitory control tasks (e.g., Johnstone et al., 2007).

TRAINING EF IN HEALTHY INDIVIDUALS

Given that EFs are subject to significant developmental progress across childhood and up to late adolescence and because they are significant predictors for many life outcomes (see Introduction), numerous studies aimed at improving these control functions by means of cognitive training interventions. Even though most of these studies have been restricted to the assessment of cognitive abilities via experimental training and transfer tasks at the lab, their ultimate goal is to improve children's EF in order to facilitate typical activities in their daily lives, such as learning and academic development (for a review, see Titz and Karbach, 2014). Even though evidence for the transfer of EF training to activities of daily living is still limited (see Potential for the Application in Clinical and Educational Settings), the existing studies have provided important insights into the mechanisms underlying behavioral plasticity (see Training EF in Healthy Individuals) and their neural underpinnings. In keeping with Miyake et al.'s (2000) model of EF, we selectively review studies that have trained cognitive flexibility/shifting, working memory/updating, and inhibition in school-aged children and adolescents.

COGNITIVE FLEXIBILITY/SHIFTING

While studies investigating cognitive flexibility in preschoolers have often applied card sorting tasks such as the DCCS (Zelazo, 2006), training studies including children over the age of 6 years have mostly relied on task-switching training in order to improve cognitive flexibility. In task-switching studies, participants are instructed to perform two or more simple decision tasks and to switch between them upon a specific cue or in a specific order. For instance, they may be required to decide whether a picture presented on the computer screen shows a fruit or a vegetable (task A) on some trials and to decide whether the picture is small or large (task B) on other trials (cf. Karbach and Kray, 2009). Comparing performances in task-homogeneous blocks (only task A or task B has to be performed) to performances in task-heterogeneous blocks (participants have to switch between task A and B) allows assessing the ability to maintain and select two task sets (general switch costs). Comparing the performances on switch trials (AB, BA) to performances on stay trials (AA, BB) provides a measure for cognitive flexibility (specific switch costs; for a review, see Monsell, 2003).

While previous studies have shown that task maintenance and selection as well as cognitive flexibility improved after task-switching training in children and adolescents between the ages of 7 and 20 years (e.g., Cepeda et al., 2001; Kray et al., 2008, 2013; Karbach and Kray, 2009; Zinke et al., 2012), there is also evidence for transfer of task-switching training to new untrained tasks. When it comes to childhood, a study including a sample of children between the ages of 7 and 9 years as well as younger adults and older adults ($N = 168$), investigated the effects of four sessions of intensive internally cued task-switching training. Compared to a control condition including the same tasks without the switching component (single-task training), the task-switching training resulted in performance improvements on a structurally similar untrained switching task (near transfer), as well as measures of inhibition, verbal and visuo-spatial working memory and reasoning (far transfer) (Karbach and Kray, 2009; see also Kray et al., 2012). The findings of this study showed that the transfer of task-switching training was not merely mediated by an automatization of the two component tasks A and B (which were also trained in the control condition). Moreover, a comparison of different training conditions showed that children's transfer was reduced when the training tasks were different switching tasks in each one of the four training sessions (as compared to the same task in each session), while the opposite pattern was found in adults. Thus, the increased cognitive load associated with the need to adapt to new training tasks in each session may not have left enough processing capacity to implement the trained abilities and to develop cognitive representations of the task structure (cf. van Merriënboer et al., 2006). One interesting finding was the broad scope of transfer, especially considering that transfer in many other training studies was more limited (see Shipstead et al., 2012; Melby-Lervåg and Hulme, 2013; Redick et al., 2013). The authors attributed the transfer to the nature of the training task, which required a number of EF abilities: demands on goal maintenance and task set-selection were high during training, because participants had to maintain the task sequence, as they did not receive external task cues. Moreover, inhibitory control was required at all times because the stimuli were ambiguous (i.e., they always represented features relevant for both component tasks). Therefore, the training of multiple EF abilities may have supported transfer to other executive and cognitive tasks.

A similar training regime was investigated in a sample of adolescents (10–14 years of age, $N = 80$). In this study, the task-switching training was compared to a passive control group, a physical exercise group and a group performing task-switching training and physical training (Zinke et al., 2012). Analyses showed a reduction of specific switch costs over the course of the training. These improvements were driven by larger training-related benefits on switch trials as compared to stay trials, suggesting that

the training specifically benefited the ability to switch between tasks and not merely increased the general speed of response execution. Interestingly, three sessions of task-switching training yielded performance benefits on an untrained switching task (near transfer) and in terms of choice reaction time and updating, but not inhibition (far transfer), without additional benefits of the acute boosts of exercise before training. Thus, these results based on adolescent participants were generally consistent with those from children and adults, indicating that task-switching training improved cognitive flexibility and transferred to tasks assessing other dimensions of EF, particularly working memory. The fact that transfer in adolescents was less pronounced than in other age groups may be due to small changes in the training regimen (e.g., the reduction of training sessions), but it may also reflect different developmental trajectories in adolescence (cf. Huizinga et al., 2006) rendering individuals more or less amenable to the effects of cognitive training than other age groups (Zinke et al., 2012).

WORKING MEMORY/UPDATING

For the assessment of children, one prototypical working memory task is the single *n*-back task, including the presentation of sequences of stimuli, such as digits or pictures. Participants are instructed to respond if the current stimulus matches the one presented *n* trials earlier in the sequence (e.g., Jonides and Smith, 1997). Another frequently applied type of task to assess working memory is the working memory span task. The simple version assesses the maximum capacity of items that can be held in working memory, for instance by instructing participants to remember sequences of digits or spatial positions. In complex working memory span tasks, this memory task has to be performed against a background processing task, such as counting or reading (e.g., Oberauer et al., 2003; Kane et al., 2004).

Given that working memory deficits are associated with a number of developmental disorders and learning difficulties, such as ADHD, dyslexia, and dyscalculia (e.g., Barkley, 1997; Schuchardt et al., 2008) many training studies have focused on clinical samples (such as, for instance, CogMed training; see Potential for the Application in Clinical and Educational Settings) and studies assessing the effects of process-based working memory training on healthy children and adolescents are surprisingly scarce. Jaeggi et al. (2011) assessed the effects of 4 weeks (≈19 sessions) of visuo-spatial *n*-back training in a sample of elementary and middle school students (mean age = 9 years, $N = 62$). Compared to a control group that performed knowledge and vocabulary-based tasks, the working-memory training group showed significant improvement on the training tasks, but no transfer to measures of fluid intelligence. Interestingly, a comparison between the participants that showed large improvements on the training task and those that only displayed small benefits revealed a differential pattern of results regarding the transfer gains: improvements on the training tasks were positively correlated with transfer gains in terms of fluid intelligence, and only the group with large improvements on the training tasks showed significant transfer of the working-memory training to measures of fluid intelligence (matrix reasoning). These differences were found immediately after training as well as three months later. The authors concluded that transfer of working memory training to fluid intelligence depends on participant's individual improvement on the training task. Moreover, they found that the perceived difficulty of the training task negatively affected training-related gains, pointing to the importance of adaptive task difficulties that are optimally challenging at all times (Jaeggi et al., 2011).

Complex working memory span tasks have been applied in two recent training studies (Loosli et al., 2012; Karbach et al., 2014). Both studies included adaptive training on child friendly tasks from the Braintwister working-memory training battery (Buschkuehl et al., 2008). In these tasks, participants were to remember the sequence of animal pictures (memory task) while analyzing the orientation of each presented picture (processing task), so that successful task execution mainly relied on verbal working memory processes. One of the studies compared the training to a passive control condition (Loosli et al., 2012; $N = 40$, 9–11 years of age; 10 sessions of training) and the other one to an active control group performing a non-adaptive low-level version of the training tasks (Karbach et al., 2014; $N = 28$, 7–9 years of age; 14 sessions of training). The results of both studies were very consistent with respect to reliable performance improvements on the training tasks in the training group and in terms of far transfer to reading abilities (for details, see Potential for the Application in Clinical and Educational Settings). Despite near transfer to a new, untrained working memory task (Karbach et al., 2014), no further transfer to any other experimental tasks occurred across studies, including measures of cognitive flexibility, inhibition, and fluid intelligence. Compared with Jaeggi et al.'s (2011) findings, this data suggests that visuo-spatial working-memory training might be more effective in order to induce transfer to other domains of EF and other cognitive abilities than verbal working-memory training. One line of evidence supporting this idea is the literature on the broad transfer of CogMed working-memory training, which includes both verbal and visuo-spatial training tasks (Klingberg, 2010; for details, see Potential for the Application in Clinical and Educational Settings). However, a systematic comparison of verbal and visuo-spatial working-memory training and the subsequent transfer effects is needed to test this hypothesis.

Moreover, transfer of working-memory training fits well with recent findings from neuroimaging studies on adults. Training on updating and switching tasks (Dahlin et al., 2008; Karbach and Brieber, 2010) has been shown to reduce activity in fronto-parietal networks and increase activity in the striatum (see also Olesen et al., 2004), a structure that is of particular importance for learning processes. It serves as a gating mechanism that decides which processes need to be worked on by the frontal and parietal areas of the brain. Thus, this increased activity in the striatum and, at the same time, the decreased fronto-parietal activation may be indicative of more automated task processing after the training and may suggest a shifted from a broad, dispersed network to a specific and more optimal one mediating efficient executive control processes.

In sum, recent findings from childhood and early adolescence showed that working-memory training has the potential to improve both verbal and visuo-spatial working memory ability (i.e., performance gains on the trained tasks). Evidence for

near and especially far transfer of working-memory training to untrained tasks and abilities has been reported less consistently and has been discussed very controversially (e.g., Shipstead et al., 2012; Redick et al., 2013). It seems to depend on the nature of the training, the transfer tasks, and the control condition as well as the baseline performance and the motivation of the participants (e.g., Jaeggi et al., 2011, 2014; Shah et al., 2012; Green et al., 2013; Titz and Karbach, 2014). Results from studies on adults suggest, for instance, that the updating component of working memory (e.g., storage and processing) has to be engaged during training in order to induce transfer to untrained working memory tasks and reasoning (von Bastian and Oberauer, 2013; see also Zinke et al., 2014).

INHIBITION

A widely used inhibition task is the Stroop task (Stroop, 1935), that requires participants to respond to the font color of words. The stimuli are either congruent (e.g., GREEN in green font color) or incongruent (e.g., GREEN in blue font color). Responses to incongruent stimuli are usually slower and more erroneous than to congruent stimuli (Stroop effect) reflecting the cognitive effort associated with the need to overcome the tendency to produce the more automated action of reading the word instead of naming the font color. Interference control is often assessed by means of the Flanker task (Eriksen and Eriksen, 1974) that requires participants to respond to a stimulus that is flanked by two other stimuli on each side. These stimuli can be congruent (e.g., HHHHH) or incongruent (e.g., SSHSS). Again, responses to incongruent stimuli are typically slower and more erroneous than to congruent stimuli (Flanker effect) reflecting the difficulty of focusing on the stimulus in the middle while suppressing interference from the surrounding letters.

However, even though it has been reported that inhibitory skills could be trained in preschoolers (e.g., Thorell et al., 2009), studies incorporating typical inhibition tasks into their training regimes for older children are rare. To our knowledge, there is no report of a training exclusively relying on inhibition tasks. Still, the training battery applied by Rueda et al. (2005) included a number of tasks tapping stimulus discrimination, conflict resolution, and inhibition, but also visual attention and anticipation exercises. The authors investigated the effects of five sessions of this "executive attention" training in 4- and 6-year-olds. After training, there was no significant near transfer to interference control (measured by the Attention Network Test, a child-friendly version of the Flanker paradigm) but far transfer to intelligence test scores, especially on the matrices scale. Thus, the training battery did not benefit EF, but reasoning abilities in preschoolers and first graders.

Even though explicit inhibition trainings are so scarce, it should be noted that trainings from the domain of working memory and cognitive flexibility often implicitly trained a fair amount of interference control. The task switching studies described above (Karbach and Kray, 2009; Zinke et al., 2012), for instance, included ambiguous stimuli (i.e., stimuli representing features that were relevant for both tasks, such as a large fruit or a small vegetable, for instance) and therefore the need to suppress interference from the currently irrelevant dimension (e.g., "large" when the currently relevant task was to decide between fruit or vegetable) and to focus on the relevant dimension. Moreover, the complex working memory tasks applied in other studies (e.g., Loosli et al., 2012; Karbach et al., 2014) included high demands on inhibitory control because participants had to inhibit the responses from the concurrent processing task in order to properly focus on the memory task: for instance, in the Braintwister trainings tasks, the children were to ignore their responses regarding the orientation of the animals (processing task) and to focus or their sequence (memory task).

Whether and to what extent inhibitory abilities may be improved in older children and adolescents remains to be examined. Future studies may rely on inhibition tasks, such as the Stroop-task, the stop-signal task, or antisaccade tasks, or on interference control tasks, such as the Flanker task. Given that recent studies showed that inhibitory control in adolescents may be improved considerably by motivational factors, such as performance-related rewards (Kohls et al., 2009; Geier et al., 2010), the room for improvement may be substantial.

Thus, recent findings from process-based EF training indicate that each one of the key domains of EF can be improved by cognitive training in childhood and adolescence. There is also evidence for transfer of EF training to other dimensions of EF (e.g., from task-switching training to working memory abilities), supporting the view that executive control is a multifaceted construct including a number of correlated but separable control dimensions (e.g., Miyake et al., 2000; Miyake and Friedman, 2012). The fact that EF training also benefited performance on fluid intelligence tasks (especially matrix reasoning) is consistent with recent latent variable approaches that confirmed a strong relationship between both domains in childhood and adulthood (e.g., Friedman et al., 2006; Engel de Abreu et al., 2010). Future studies need to shed light on specific features of the training regimes and characteristics of the participants that have the potential to support positive effects of cognitive training interventions. This will probably include a shift from the general question of whether a given training in effective or not (i.e., the comparison of mean group differences) to more fine grained analyses testing individual differences in order to determine for whom the training actually works.

INDIVIDUAL DIFFERENCES IN TRAINING-INDUCED GAINS

The selected findings reviewed above, along with numerous other studies from the field of cognitive training research (for reviews, see Hertzog et al., 2008; Lustig et al., 2009; Noack et al., 2009; Karbach and Schubert, 2013; Kray and Ferdinand, 2013; Titz and Karbach, 2014), showed that cognitive training interventions have the potential to yield significant EF benefits and even transfer of EF training at the group level, but evidence on individual differences is still limited, especially in childhood and adolescence. This is particularly critical in populations displaying rapid cognitive developmental progress, because children and adolescents are likely to differ more from each other than young adults and between-group comparisons do little justice to individuals' strengths and weaknesses. Therefore, the question who benefits most from cognitive interventions has been more and more acknowledged in the field of cognitive training research lately

and an increasing number of studies have analyzed why some individuals benefited more than others. The importance of this question is obvious from an applied point of view, especially when it comes to the adaptation of training interventions to populations with specific needs, such as students with cognitive or academic deficits. Moreover, it also is of interest on the theoretical level, because individual differences in training-related benefits may help us understand the underpinnings of cognitive and neural plasticity.

Two prominent accounts haven been put forward to describe and explain individual differences in training-related performance gains: first, the magnification account (also known as Matthew effect or scissor effect) assumes that individuals that are already performing very well will also benefit most from cognitive interventions. It is assumed that high-performing and well-educated participants have more efficient cognitive resources to acquire and implement new strategies and abilities. Thus, baseline cognitive performance at pretest should be positively correlated with the training-related gains. And with respect to EF, which gradually develop across childhood and adolescence (see Age-related Changes in EF), cognitive interventions should result in a magnification of age differences and individual differences. In fact, there are a number of earlier studies supporting this account, most of them from the field of memory strategy training, for instance by means of the method of loci (e.g., Björklund and Douglas, 1997; Brehmer et al., 2007; see Verhaeghen et al., 1992 for a meta-analysis).

Second, the compensation account assumes that high-performing individuals will benefit less from cognitive interventions, because they are already functioning at the optimal level and therefore have less room for improvement. Thus, baseline cognitive performance should be negatively correlated with training gains and age differences and individual differences should be reduced after the intervention. Evidence supporting the compensation account comes from numerous studies focusing on EF training, revealing that training-related benefits were larger in children and older adults than in younger adults (e.g., Kramer et al., 1995; Kray and Lindenberger, 2000; Cepeda et al., 2001; Bherer et al., 2008; Kray et al., 2008; Karbach and Kray, 2009; Dorbath et al., 2011). While these studies were based on comparisons at the group level, recent studies also have analyzed correlations between baseline cognitive ability and training-related benefits, indicating that working memory training yielded larger training and transfer effects in children with low cognitive performance at pretest (Jaeggi et al., 2008; Dahlin, 2011; Karbach et al., 2014; but see Loosli et al., 2012; Holmes and Gathercole, 2013).

Moreover, recent work has applied latent variable approaches to analyze individual differences in performance changes as well as correlations between baseline cognitive ability and training-related benefits. One of these studies provided evidence for the magnification account: Loevdén et al. (2012) analyzed data from a study on episodic memory strategy training (based on the method of loci), including children and adolescents (9–12 years) as well as younger adults (20–25 years) and older adults (65–78 years). Even though strategy instructions at the beginning of training reduced individual differences in memory performance, further training ultimately magnified individual differences. In contrast, a study on process-based task-switching training provided evidence for the compensation account (Karbach and Spengler, 2012): children (8–10 years), younger adults (18–26 years), and older adults (62–76 years) consistently showed a reduction of age differences and individual differences not only in terms of performance gains on the training task, but also for transfer to a new, untrained switching. Moreover, cognitive performance at pretest was negatively correlated with training and transfer gains, suggesting that low-performing participants showed larger training-induced gains.

Taken together, research investigating the role of baseline cognitive ability for training-related performance gains indicated that magnification effects were more likely in the domain of strategy training, whereas compensation effects were found more often after process-based training interventions, such as EF training (for further comments on the difference between strategy-based and process-based interventions see Lustig et al., 2009; Noack et al., 2009; Kliegel and Bürki, 2012; Verhaeghen, 2014).

POTENTIAL FOR THE APPLICATION IN CLINICAL AND EDUCATIONAL SETTINGS

The cognitive and neural plasticity uncovered in the field of cognitive training research certainly has important implications for applied settings, such as clinical or educational programs. At this point, results from well-controlled training studies conducted in these areas are still limited, but the latest basic research findings may be very informative for the design of applied training programs. In order to illustrate the potential of EF training for clinical end educational settings, we briefly review selected research findings on (1) EF training in samples suffering from ADHD and (2) the effects of EF training on academic achievement in childhood and adolescence.

EFFECTS OF EF TRAINING IN PARTICIPANTS SUFFERING FROM ADHD

Attention-deficit hyperactivity disorder is typically characterized by the three behavioral core symptoms inattention, impulsivity, and hyperactivity (Barkley, 1997). Consistently, children with ADHD usually show cognitive impairments in terms of working memory, inhibitory control, and attention. Many life outcomes are negatively affected by the disorder, such as academic development, vocational success, and social interactions (cf. Shah et al., 2012). It is therefore not surprising that many cognitive training studies have aimed at compensating cognitive and behavioral symptoms and supporting the social and scholastic development of children with ADHD.

One recent study has adopted a task-switching training regime that was effective in healthy individuals. Boys between the ages of 7 and 12 years diagnosed with ADHD and medicated with methylphenidate performed four sessions of intensive task-switching training. Compared to a single-task practice condition, the task-switching training benefited inhibition and working memory, both of which are typically impaired in children with ADHD, but not fluid intelligence (Kray et al., 2012). These findings indicate that even relatively short interventions have the potential to selectively improve cognitive deficits associated with ADHD.

A number of other studies have applied the CogMed training battery that has been designed to improve working memory and

executive control[1] (see also Klingberg et al., 2005). It includes a variety of verbal and visuo-spatial short-term memory and working memory tasks that are usually trained for 25 sessions. Several studies have provided evidence for the effectiveness of CogMed training in children with ADHD (e.g., Klingberg et al., 2002, 2005). After the training, children improved their performances on new untrained working memory tasks, but also on measures of inhibition and fluid intelligence. The authors attributed these findings to increased neural efficiency in overlapping neural circuits that were recruited for performing the training and the transfer tasks (Klingberg et al., 2005). Moreover, these improvements were also observed in terms of parent-rated symptoms of inattention and hyperactivity/impulsivity and many gains were maintained for three months (e.g., Klingberg et al., 2005; see also Beck et al., 2010).

Despite these and other encouraging findings, recent reviews and meta-analyses suggest that the effects of working memory training in children with ADHD may not be that far reaching (Rapport et al., 2013; Chacko et al., 2014). While many interventions resulted in significant improvements on the training task and on structurally similar near transfer tasks, particularly far transfer to untrained cognitive domains, behavioral symptoms, and academic outcomes was not significant. Moreover, a few methodological flaws were criticized, such as the use of non-adaptive or no-contact control groups, the use of individual tasks instead of batteries in order to measure constructs, or the analysis of reports from parents that were not blind to training conditions (Shipstead et al., 2012). In fact, a well-controlled recent study on the effects of CogMed training in 7- to 11-year-old children with ADHD did not show beneficial effects beyond working memory improvements (Chacko et al., 2014; see also Holmes et al., 2009b). Thus, should we consider this type of cognitive intervention ineffective in children with ADHD? As nicely summarized by Gathercole (2014) a recent comment, the answer to this question should definitively be no. Instead, we agree that it will be crucial to overcome methodological issues by designing new approaches yielding functional transfer with relevant cognitive benefits. According to Gathercole (2014), this may for instance be achieved by designing hybrid training protocols including features of different types of training that have been proven beneficial, such as *n*-back training (e.g., Jaeggi et al., 2011) and task-switching training (Karbach and Kray, 2009; Kray et al., 2012). Moreover, she suggested to directly implement adaptive training methods into activities that children with ADHD struggle with in the classroom, such as mental arithmetic, following instructions, and language comprehension.

Another issue that has to be considered is the fact that there usually are large individual differences in the effectiveness of cognitive interventions (see Individual Differences in Training-induced Gains; Titz and Karbach, 2014). This is particularly important in children suffering from ADHD, because there is a large variety of causes for the cognitive and behavioral symptoms, such as genetics, anxiety, life stress, exposure to environmental toxins, etc. (Millichap, 2008; Shah et al., 2012). In addition, children may differ with respect to the treatments they previously received as well as regarding their motivation to comply with the training protocol. These and other factors may very well result in big individual differences in training-induced gains, and these differences may mask large individual gains if data is only analyzed on the group-level (Shah et al., 2012). Unfortunately, many clinical studies have not analyzed individual differences, most likely because the sample sizes were too small. While this certainly is an important issue to deal with in future studies, increasing the sample sizes may also come at a cost: according to Shipstead et al. (2012), many small-scale studies reported effects that were not replicated in larger-scale interventions. However, this may not necessarily indicate that the training is not effective, but may be caused by the difficulty of maintaining the integrity of the exact training protocol in a larger context: "Resource constrains may give researchers a difficult choice: small-scale studies that do not have the sample size to consider the role of individual differences, or larger-scale studies that allow one to assess individual differences but cannot have the same level of experimenter control" (Shah et al., 2012, p. 205).

In sum, more fine-grained analyses of the mechanisms underlying transfer of cognitive training and individual differences therein is clearly needed in order to determine which type of training may be most beneficial for children suffering from ADHD or other neurocognitive or developmental disorders. The present results nonetheless show considerable cognitive and neural plasticity in children suffering from ADHD (Jolles and Crone, 2012; Rapport et al., 2013; Chacko et al., 2014), indicating that the individual benefits of well-tailored cognitive interventions may be considerable.

EFFECTS OF EF TRAINING ON ACADEMIC ACHIEVEMENT

Research on academic achievement has repeatedly confirmed EF, and particularly working memory, as important prerequisites for the general ability to acquire knowledge and new skills. EF are not only related to higher-level cognitive abilities contributing to academic success, such as problem solving, but also to performance in the classroom (for a review, see Titz and Karbach, 2014). In fact, EF have been shown to explain at least as much variance in academic achievement as intelligence (e.g., Swanson, 2004; Altemeier et al., 2008; Andersson, 2008; Alloway and Alloway, 2010; Lu et al., 2011), which is usually considered the most powerful predictor of academic success (e.g., Gottfredson, 2002; cf. Gustafsson and Undheim, 1996).

Studies investigating the contribution of EF to scholastic achievement have often focused on the domains of language and mathematics and showed that EF are directly associated with math ability as well as with reading, writing, and language comprehension (Titz and Karbach, 2014). The strong association between EF and academic achievement is also supported by findings showing that children suffering from developmental disorders or learning disabilities often display specific EF deficits, suggesting these deficits are risk factors for poor academic performance and development (e.g., Barkley, 1997; Gathercole et al., 2006; Schuchardt et al., 2008; see also Effects of EF Training in Participants Suffering from ADHD). Considering this strong relation between EF and academic abilities, one may assume that even small increases in EF functioning might improve children's academic performance.

However, despite the growing number of cognitive training studies, only very few of them included transfer tasks from

[1] www.cogmed.com

the domain of academic abilities. Most of these studies have applied working memory training regimes to children with cognitive deficits or learning difficulties (for an extensive review, see Titz and Karbach, 2014). This work showed that 25 sessions of CogMed working memory training transferred to new, untrained working memory tasks in children with low working memory ability (8–11 years of age), but not to reading or mathematical reasoning abilities (Holmes et al., 2009a; Dunning et al., 2013). In contrast, a recent field study from the same group showed that teacher administered CogMed working memory training improved performance on standardized tests for English and math in sixth grade (Holmes and Gathercole, 2013), indicating that training-induced memory improvements may transfer to ecologically valid measures of academic achievement in low-achieving students. These findings are supported by results showing that students with special educational needs and attention problems (9–12 years) benefited in terms of reading comprehension and basic number skills (Dahlin, 2011, 2013). In contrast to CogMed training, an interactive working memory training game called Jungle Memory including 32 sessions of verbal and visuo-spatial working memory tasks, yielded no transfer to performance on tasks assessing arithmetic and spelling in children with learning difficulties (mean age = 10.10 years; Alloway et al., 2013).

As for healthy children, two recent studies have applied tasks from the Braintwister working memory training battery (Buschkuehl et al., 2008) that included complex verbal working memory tasks. After 10–14 sessions of training, both studies consistently showed improvements on standardized tests of reading in students between 7 and 11 years of age (Loosli et al., 2012; Karbach et al., 2014). In both studies, the authors attributed the transfer to reading to the strong relation of complex span tasks to reading ability (e.g., Daneman and Merikle, 1996; Engel de Abreu et al., 2011) and memory retrieval (Unsworth and Engle, 2006).

In sum, recent developmental findings indicated that cognitive training might indeed compensate EF deficits in children with ADHD and support school-related abilities and academic performance. However, it is also obvious that these effects are not consistent across studies and it is still unknown to what extent they may be modulated by age-related differences in social and emotional processes or by motivational components. Nevertheless, the existing findings are encouraging because they demonstrate the potential of cognitive training for improving daily life performance outside of the lab, even if much more research is needed to fully uncover the underlying mechanisms and to identify training regimes that reliably and consistently improve specific areas of academic performance and development or specific cognitive deficits in clinical samples.

CONCLUSION

A large body of research has confirmed the multidimensional structure of EF and the importance of fronto-parietal networks for the integrity and development of executive control. Considering the contribution of EF to various life outcomes, many studies have investigated the effectiveness of cognitive training interventions designed to improve EF. This research showed that cognitive plasticity is considerable across the lifespan, even up to very old age. It has also been suggested that behavioral and neural plasticity are especially high in childhood and the prefrontal lobes are particularly sensitive to environmental influences in that age group. Consistently, research on children and adolescents showed that process-based EF training is an effective means to improve control abilities, particularly working memory and cognitive flexibility. Moreover, many EF trainings benefited performance on tasks that were not trained, such as measures of attention or fluid intelligence, even though other studies suggested that these effects may neither be robust nor consistent. Recent work suggest that they may be (a) increased if the training and the transfer task share overlapping processing components and brain regions and (b) more likely after process-based trainings than after interventions teaching task-specific strategies.

The analysis of individual differences in training-induced gains showed that process-based interventions have yielded compensation effects with larger gains in participants that scored worse at pretest. These findings suggest that process-based trainings may be particularly useful for compensating specific EF deficits associated with neurodevelopmental disorders and learning difficulties. In fact, earlier research on working-memory training and task-switching training resulted in significant training gains and broad transfer effects in children with ADHD, even though recent studies have challenged these optimistic results to a certain degree.

Aside from clinical settings, recent studies have also focused on educational contexts. The few existing studies have provided mixed but encouraging findings, indicating that working-memory training has the potential to improve academic abilities, particularly in the domain of language and reading. These benefits have not only been reported for normally developing children, but also for students with cognitive deficits and learning difficulties. Cleary, further research is needed to improve the understanding of the mechanisms mediating transfer of cognitive training to academic abilities. These studies will be of major importance for tailoring training interventions to the specific needs of certain populations or individuals. Moreover, future studies may want to assess how social and emotional development is related to training-induced improvements and to which degree training-related benefits may be driven by motivational components.

REFERENCES

Alloway, T., and Alloway, R. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *J. Exp. Child Psychol.* 106, 20–29. doi: 10.1016/j.jecp.2009.11.003

Alloway, T., Bibile, V., and Lau, G. (2013). Computerized working memory training: can it lead to gains in cognitive skills in students? *Comput. Hum. Behav.* 29, 632–638. doi: 10.1016/j.chb.2012.10.023

Altamirano, L. J., Miyake, A., and Whitmer, A. J. (2010). When mental inflexibility facilitates executive control beneficial side effects of ruminative tendencies on goal maintenance. *Psychol. Sci.* 21, 1377–1382. doi: 10.1177/0956797610381505

Altemeier, L. E., Abbott, R. D., and Berninger, V. W. (2008). Executive functions for reading and writing in typical literacy development and dyslexia. *J. Clin. Exp. Neuropsychol.* 30, 588–606. doi: 10.1080/13803390701562818

Andersson, U. (2008). Working memory as a predictor of written arithmetical skills in children: the importance of central executive functions. *Br. J. Educ. Psychol.* 78, 181–203. doi: 10.1348/000709907X209854

Ardila, A., Pineda, D., and Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Arch. Clin. Neuropsychol.* 15, 31–36. doi: 10.1093/arclin/15.1.31

Baddeley, A. D. (1986). *Working Memory*. New York: Oxford University Press.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* 121, 65. doi: 10.1037/0033-2909.121.1.65

Basak, C., Boot, W. R., Voss, M. W., and Kramer, A. F. (2008). Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychol. Aging* 23, 765. doi: 10.1037/a0013494

Beck, S. J., Hanson, C. A., Puffenberger, S. S., Benninger, K. L., and Benninger, W. B. (2010). A controlled trial of working memory training for children and adolescents with ADHD. *J. Clin. Child Adolesc. Psychol.* 39, 825–836. doi: 10.1080/15374416.2010.517162

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K., and Becic, E. (2008). Transfer effects in task-set cost and dual-task cost after dual-task training in older and younger adults: further evidence for cognitive plasticity in attentional control in late adulthood. *Exp. Aging Res.* 34, 188–219. doi: 10.1080/03610730802070068

Bjorklund, D. F. (2005). *Children's Thinking: Cognitive Development and Individual Differences*. Belmont, CA: Wadsworth.

Björklund, D. F., and Douglas, R. N. (1997). "The development of memory strategies," in *The Development of Memory in Childhood*, eds N. Cowan and C. Hulme (Sussex: Psychology Press), 201–246.

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Brehmer, Y., Li, S. C., Mueller, V., von Oertzen, T. V., and Lindenberger, U. (2007). Memory plasticity across the life span: uncovering children's latent potential. *Dev. Psychol.* 43, 465–477. doi: 10.1037/0012-1649.43.2.465

Buitenweg, J. I., Murre, J. M., and Ridderinkhof, K. R. (2012). Brain training in progress: a review of trainability in healthy seniors. *Front. Hum. Neurosci.* 6:183. doi: 10.3389/fnhum.2012.00183

Bull, R., Andrews Espy, K., Wiebe, S. A., Sheffield, T. D., and Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: sources of variation in emergent mathematic achievement. *Dev. Sci.* 14, 679–692. doi: 10.1111/j.1467-7687.2010.01012.x

Bunge, S. A. (2004). How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cogn. Affect. Behav. Neurosci.* 4, 564–579. doi: 10.3758/CABN.4.4.564

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., and Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron* 33, 1–20. doi: 10.1016/S0896-6273(01)00583-9

Bunge, S. A., and Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Curr. Opin. Neurobiol.* 17, 243–250. doi: 10.1016/j.conb.2007.02.005

Bunge, S. A., and Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Curr. Dir. Psychol. Sci.* 15, 118–121. doi: 10.1111/j.0963-7214.2006.00419.x

Buschkuehl, M., Jaeggi, S. M., and Jonides, J. (2012). Neuronal effects following working memory training. *Dev. Cogn. Neurosci.* 2, S167–S179. doi: 10.1016/j.dcn.2011.10.001

Buschkuehl, M., Jaeggi, S., Kobel, A., and Perrig, W. J. (2008). *Braintwister – Aufgabensammlung für kognitives Training*. Bern: Institut für Psychologie, Universität Bern.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., and Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, i–vi, 1–143. doi: 10.2307/1166214

Casey, B. J., Tottenham, N., Liston, C., and Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends Cogn. Sci.* 9, 104–110. doi: 10.1016/j.tics.2005.01.011

Cepeda, N. J., Kramer, A. F., and De Sather, J. C. M. G. (2001). Changes in executive control across the life span: examination of task-switching performance. *Dev. Psychol.* 37, 715–729. doi: 10.1037/0012-1649.37.5.715

Chacko, A., Bedard, A. C., Marks, D. J., Feirsen, N., Uderman, J. Z., Chimiklis, A., et al. (2014). A randomized clinical trial of Cogmed Working Memory Training in school-age children with ADHD: a replication in a diverse sample using a control condition. *J. Child Psychol. Psychiatry* 55, 247–255. doi: 10.1111/jcpp.12146

Chevalier, N., and Blaye, A. (2009). Setting goals to switch between tasks: effect of cue transparency on children's cognitive flexibility. *Dev. Psychol.* 45, 782. doi: 10.1037/a0015409

Collette, F., Hogge, M., Salmon, E., and Van der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience* 139, 209–221. doi: 10.1016/j.neuroscience.2005.05.035

Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., et al. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Hum. Brain Mapp.* 25, 409–423. doi: 10.1002/hbm.20118

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755

Crone, E. A., Bunge, S. A., Van der Molen, M. W., and Ridderinkhof, K. R. (2006). Switching between tasks and responses: a developmental study. *Dev. Sci.* 9, 278–287. doi: 10.1111/j.1467-7687.2006.00490.x

Crone, E. A., Ridderinkhof, K. R., Worm, M., Somsen, R. J. M., and Van Der Molen, M. W. (2004). Switching between spatial stimulus–response mappings: a developmental study of cognitive flexibility. *Dev. Sci.* 7, 443–455. doi: 10.1111/j.1467-7687.2004.00365.x

Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., and Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science* 320, 1510–1512. doi: 10.1126/science.1155466

Dahlin, K. (2013). Working memory training and the effect on mathematical achievement in children with attention deficits and special needs. *J. Educ. Learn.* 2, 118–133. doi: 10.5539/jel.v2n1p118

Dahlin, K. I. E. (2011). Effects of working memory training on reading in children with special needs. *Read. Writ.* 24, 479–491. doi: 10.1007/s11145-010-9238-y

Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546

Davidson, M. C., Amso, D., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006

Dawson, P., and Guare, R. (2010). *Executive Skills in Children and Adolescents: A Practical Guide to Assessment and Intervention*. New York: Guilford Press.

de Frias, C. M., Dixon, R. A., and Strauss, E. (2006). Structure of four executive functioning tests in healthy older adults. *Neuropsychology* 20, 206. doi: 10.1037/0894-4105.20.2.206

Denckla, M. B., and Reiss, A. L. (1997). "Prefrontal-subcortical circuits in developmental disorders," in *Development of the Prefrontal Cortex: Evolution, Neurobiology, and Behavior*, eds N. A. Krasnegor, G. R. Lyon, and P. S. Goldman-Rakic (Baltimore, MD: Paul H. Brooks), 283–293.

Diamond, A. (2006). "The early development of executive functions," in *Lifespan Cognition: Mechanisms of Change*, eds E. Bialystok and F. Craik (New York, NY: Oxford University Press), 70–95. doi: 10.1093/acprof:oso/9780195169539.003.0006

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Curr. Dir. Psychol. Sci.* 21, 335–341. doi: 10.1177/0963721412453722

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Dorbath, L., Hasselhorn, M., and Titz, C. (2011). Aging and executive functioning: a training study on focus-switching. *Front. Psychol.* 2:257. doi: 10.3389/fpsyg.2011.00257

Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., and Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends Cogn. Sci.* 12, 99–105. doi: 10.1016/j.tics.2008.01.001

Duncan, J., Emslie, H., Williams, P., Johnson, R., and Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn. Psychol.* 30, 257–303. doi: 10.1006/cogp.1996.0008

Duncan, J., and Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–483. doi: 10.1016/S0166-2236(00)01633-7

Dunning, D. L., Holmes, J., and Gathercole, S. E. (2013). Does working memory training lead to generalized improvements in children with low working memory? A randomized controlled trial. *Dev. Sci.* 16, 915–925. doi: 10.1111/desc.12068

Edin, F., Macoveanu, J., Olesen, P., Tegnér, J., and Klingberg, T. (2007). Stronger synaptic connectivity as a mechanism behind development of working memory-related brain activity during childhood. *J. Cogn. Neurosci.* 19, 750–760. doi: 10.1162/jocn.2007.19.5.750

Eigsti, I. M., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M., et al. (2006). Predicting cognitive control from preschool to late adolescence and young adulthood. *Psychol. Sci.* 17, 478–484. doi: 10.1111/j.1467-9280.2006.01732.x

Engel de Abreu, P. M. J., Conway, A. R. A., and Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence* 38, 552–561. doi: 10.1016/j.intell.2010.07.003

Engel de Abreu, P. M. J., Gathercole, S. E., and Martin, R. (2011). Disentangling the relationship between working memory and language: the roles of short-term storage and cognitive control. *Learn. Individ. Differ.* 21, 569–574. doi: 10.1016/j.lindif.2011.06.002

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Fair, D. A., Cohen, A. L., Dosenbach, N. U., Church, J. A., Miezin, F. M., Barch, D. M., et al. (2008). The maturing architecture of the brain's default network. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4028–4032. doi: 10.1073/pnas.0800376105

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., and Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychol. Sci.* 17, 172–179. doi: 10.1111/j.1467-9280.2006.01681.x

Friedman, N. P., Miyake, A., Robinson, J. L., and Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: a behavioral genetic analysis. *Dev. Psychol.* 47, 1410. doi: 10.1037/a0023750

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *J. Exp. Psychol. Gen.* 137, 201. doi: 10.1037/0096-3445.137.2.201

Gathercole, S., Alloway, T., Willis, C., and Adams, A. (2006). Working memory in children with reading disabilities. *J. Exp. Child Psychol.* 93, 265–281. doi: 10.1016/j.jecp.2005.08.003

Gathercole, S. E. (2014). Commentary: working memory training and ADHD – where does its potential lie? Reflections on Chacko et al. (2013). *J. Child Psychol. Psychiatry* 55, 256–257. doi: 10.1111/jcpp.12196

Gathercole, S. E., Pickering, S. J., Ambridge, B., and Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* 40, 177. doi: 10.1037/0012-1649.40.2.177

Geier, C. F., Terwilliger, R., Teslovich, T., Velanova, K., and Luna, B. (2010). Immaturities in reward processing and its influence on inhibitory control in adolescence. *Cereb. Cortex* 20, 1613–1629. doi: 10.1093/cercor/bhp225

Gogtay, N., Giedd, J. N., Hayashi, K. M., Greenstein, D., Vaituzis, C., Nugent, T. F., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179. doi: 10.1073/pnas.0402680101

Gottfredson, L. S. (2002). Where and why g matters: not a mystery. *Hum. Perform.* 15, 25–46. doi: 10.1080/08959285.2002.9668082

Green, C., Strobach, T., and Schubert, T. (2013). On methodological standards in training and transfer experiments. *Psychol. Res.* doi: 10.1007/s00426-013-0535-3 [Epub ahead of print].

Gruber, A. J., Dayan, P., Gutkin, B. S., and Solla, S. A. (2006). Dopamine modulation in the basal ganglia locks the gate to working memory. *J. Comput. Neurosci.* 20, 153–166. doi: 10.1007/s10827-005-5705-x

Gustafsson, J.-E., and Undheim, J. O. (1996). "Individual differences in cognitive functions," in *Handbook of Educational Psychology*, eds D. C. Berliner and R. C. Calfee (New York: Prentice Hall International), 186–242.

Hertzog, C., Kramer, A. F., Wilson, R. S., and Lindenberger, U. (2008). Enrichment effects on adult cognitive development can the functional capacity of older adults be preserved and enhanced? *Psychol. Sci. Public Interest* 9, 1–65.

Heyder, K., Suchan, B., and Daum, I. (2004). Cortico-subcortical contributions to executive control. *Acta Psychol.* 115, 271–289. doi: 10.1016/j.actpsy.2003.12.010

Holmes, J., and Gathercole, S. E. (2013). Taking working memory training from the laboratory into schools. *Educ. Psychol.* doi: 10.1080/01443410.2013.797338 [Epub ahead of print].

Holmes, J., Gathercole, S. E., and Dunning, D. L. (2009a). Adaptive training leads to sustained enhancement of poor working memory in children. *Dev. Sci.* 12, F9–F15. doi: 10.1111/j.1467-7687.2009.00848.x

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., and Elliott, J. (2009b). Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. *Appl. Cogn. Psychol.* 24, 827–836. doi: 10.1002/acp.1589

Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Dev. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Huizinga, M., and van der Molen, M. W. (2007). Age-group differences in set-switching and set-maintenance on the Wisconsin Card Sorting Task. *Dev. Neuropsychol.* 31, 193–215. doi: 10.1080/87565640701190817

Jaeggi, S., Buschkuehl, M., Jonides, J., and Perrig, W. (2008). Improving intelligence by training working memory. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6829–6833. doi: 10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Jonides, J., and Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10081–10086. doi: 10.1073/pnas.1103228108

Jaeggi, S. M., Buschkuehl, M., Shah, P., and Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Mem. Cogn.* 42, 464–480. doi: 10.3758/s13421-013-0364-z

Johnstone, S. J., Dimoska, A., Smith, J. L., Barry, R. J., Pleffer, C. B., Chiswick, D., et al. (2007). The development of stop-signal and Go/Nogo response inhibition in children aged 7-12 years: performance and event-related potential indices. *Int. J. Psychophysiol.* 63, 25–38. doi: 10.1016/j.ijpsycho.2006.07.001

Jolles, D. D., and Crone, E. A. (2012). Training the developing brain: a neurocognitive perspective. *Front. Hum. Neurosci.* 6:76. doi: 10.3389/fnhum.2012.00076

Jonides, J., and Smith, E. E. (1997). "The architecture of working memory," in *Cognitive Neuroscience*, ed. M. D. Rugg (Cambridge, MA: The MIT Press), 243–276.

Jurado, M. B., and Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychol. Rev.* 17, 213–233. doi: 10.1007/s11065-007-9040-z

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., and Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *J. Exp. Psychol. Gen.* 133, 189. doi: 10.1037/0096-3445.133.2.189

Karbach, J., and Brieber, S. (2010). "Neuronale Korrelate exekutiven Kontrolltrainings," in *Kongress der Deutschen Gesellschaft für Psychologie*, eds F. Petermann and U. Koglin (Lengerich: Pabst), 135.

Karbach, J., and Kray, J. (2007). Developmental changes in switching between mental task sets: the influence of verbal labeling in childhood. *J. Cogn. Dev.* 8, 205–236. doi: 10.1080/15248370701202430

Karbach, J., and Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Dev. Sci.* 12, 978–990. doi: 10.1111/j.1467-7687.2009.00846.x

Karbach, J., and Schubert, T. (2013). Training-induced cognitive and neural plasticity. *Front. Hum. Neurosci.* 7:48. doi: 10.3389/978-2-88919-145-1

Karbach, J., and Spengler, M. (2012). "Who benefits the most? Individual differences in the transfer of executive control training in younger and older adults," in *Abstracts of the 54th Meeting for Experimental Psychologists (TEAP)*, eds A. Bröder, E. Erdfelder, B. E. Hilbig, T. Meiser, R. F. Pohl, and D. Stahlberg (Lengerich: Pabst), 64.

Karbach, J., Strobach, T., and Schubert, T. (2014). Adaptive working-memory training benefits reading, but not mathematics in middle childhood. *Child Neuropsychol.* doi: 10.1080/09297049.2014.899336 [Epub ahead of print].

Kimberg, D. Y., D'Esposito, M., and Farah, M. J. (1997). Effects of bromocriptine on human subjects depend on working memory capacity. *Neuroreport* 8, 3581–3585. doi: 10.1097/00001756-199711100-00032

Kliegel, M., and Bürki, C. (2012). Memory training interventions require a tailor-made approach: commentary on McDaniel and Bugg. *J. Appl. Res. Mem. Cogn.* 1, 58–60. doi: 10.1016/j.jarmac.2012.01.002

Klingberg, T. (2010). Training and plasticity of working memory. *Trends Cogn. Sci.* 14, 317–324. doi: 10.1016/j.tics.2010.05.002

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., et al. (2005). Computerized training of working memory in children with ADHD – a randomized, controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* 44, 177–186. doi: 10.1097/00004583-200502000-00010

Klingberg, T., Forssberg, H., and Westerberg, H. (2002). Training of working memory in children with ADHD. *J. Clin. Exp. Neuropsychol.* 24, 781–791. doi: 10.1076/jcen.24.6.781.8395

Kohls, G., Peltzer, J., Herpertz-Dahlmann, B., and Konrad, K. (2009). Differential effects of social and non-social reward on response inhibition in children and adolescents. *Dev. Sci.* 12, 614–625. doi: 10.1111/j.1467-7687.2009.00816.x

Kramer, A. F., Larish, J. F., and Strayer, D. L. (1995). Training for attentional control in dual task settings: a comparison of young and old adults. *J. Exp. Psychol. Appl.* 1, 50–76. doi: 10.1037/1076-898X.1.1.50

Kray, J., Eber, J., and Karbach, J. (2008). Verbal self-instructions in task switching: a compensatory tool for action-control deficits in childhood and old age? *Dev. Sci.* 11, 223–236. doi: 10.1111/j.1467-7687.2008.00673.x

Kray, J., Eber, J., and Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: the role of verbalization in task preparation. *Acta Psychol.* 115, 143–165. doi: 10.1016/j.actpsy.2003.12.001

Kray, J., and Ferdinand, N. K. (2013). How to improve cognitive control in development during childhood: potentials and limits of cognitive interventions. *Child Dev. Perspect.* 7, 121–125. doi: 10.1111/cdep.12027

Kray, J., Gaspard, H., Karbach, J., and Blaye, A. (2013). Developmental changes in using verbal self-cueing in task-switching situations: the impact of task practice and task-sequencing demands. *Front. Psychol.* 4:940. doi: 10.3389/fpsyg.2013.00940

Kray, J., Karbach, J., and Blaye, A. (2012). The influence of stimulus-set size on developmental changes in cognitive control and conflict adaptation. *Acta Psychol.* 140, 119–128. doi: 10.1016/j.actpsy.2012.03.005

Kray, J., Karbach, J., Haenig, S., and Freitag, C. (2012). Can task-switching training enhance executive control functioning in children with attention deficit/-hyperactivity disorder? *Front. Hum. Neurosci.* 5:180. doi: 10.3389/fnhum.2011.00180

Kray, J., Kipp, K., and Karbach, J. (2009). The development of selective inhibitory control: the influence of verbal labeling. *Acta Psychol.* 130, 48–57. doi: 10.1016/j.actpsy.2008.10.006

Kray, J., and Lindenberger, U. (2000). Adult age differences in task switching. *Psychol. Aging* 15, 126–147. doi: 10.1037/0882-7974.15.1.126

Lehto, J. E., Juujärvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Loevdén, M., Brehmer, Y., Li, S. C., and Lindenberger, U. (2012). Training-induced compensation versus magnification of individual differences in memory performance. *Front. Hum. Neurosci.* 6:141. doi: 10.3389/fnhum.2012.00141

Loosli, S. V., Buschkuehl, M., Perrig, W. J., and Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychol.* 18, 62–78. doi: 10.1080/09297049.2011.575772

Lu, L., Weber, H. S., Spinath, F. M., and Shi, J. (2011). Predicting school achievement from cognitive and non-cognitive variables in a Chinese sample of elementary school children. *Intelligence* 39, 130–140. doi: 10.1016/j.intell.2011.02.002

Luciana, M., Conklin, H. M., Hooper, C. J., and Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Dev.* 76, 697–712. doi: 10.1111/j.1467-8624.2005.00872.x

Luna, B., Padmanabhan, A., and O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain Cogn.* 72, 101–113. doi: 10.1016/j.bandc.2009.08.005

Lustig, C., Shah, P., Seidler, R., and Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: a review and future directions. *Neuropsychol. Rev.* 19, 504–522. doi: 10.1007/s11065-009-9119-9

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1423–1442. doi: 10.1037/0278-7393.22.6.1423

Melby-Lervåg, M., and Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Dev. Psychol.* 49, 270. doi: 10.1037/a0028228

Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167

Millichap, J. G. (2008). Etiologic classification of attention-deficit/hyperactivity disorder. *Pediatrics* 121, e358–e365. doi: 10.1542/peds.2007-1332

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108

Monsell, S. (2003). Task switching. *Trends Cogn. Sci.* 7, 134–140. doi: 10.1016/S1364-6613(03)00028-7

Moriguchi, Y., and Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6017–6021. doi: 10.1073/pnas.0809747106

Moriguchi, Y., and Hiraki, K. (2011). Longitudinal development of prefrontal function during early childhood. *Dev. Cogn. Neurosci.* 1, 153–162. doi: 10.1016/j.dcn.2010.12.004

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., and O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends Cogn. Sci.* 15, 453–459. doi: 10.1016/j.tics.2011.07.011

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., and Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn. Affect. Behav. Neurosci.* 12, 241–268. doi: 10.3758/s13415-011-0083-5

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychol. Bull.* 126, 220–246. doi: 10.1037/0033-2909.126.2.220

Noack, H., Lövdèn, M., Schmiedek, F., and Lindenberger, U. (2009). Cognitive plasticity in adulthood and old age: gauging the generality of cognitive intervention effects. *Restor. Neurol. Neurosci.* 27, 435–453. doi: 10.3233/RNN-2009-0496

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behavior," in *Consciousness and Self-regulation*, ed. R. J. Davidson (New York: Plenum Press), 1–18. doi: 10.1007/978-1-4757-0629-1_1

Oberauer, K., Süß, H., Wilhelm, O., and Wittmann, W. W. (2003). The multiple faces of working memory: storage, processing, supervision, and coordination. *Intelligence* 31, 167–193. doi: 10.1016/S0160-2896(02)00115-0

Olesen, P. J., Westerberg, H., and Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nat. Neurosci.* 7, 75–79. doi: 10.1038/nn1165

O'Reilly, R. C., and Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput.* 18, 283–328. doi: 10.1162/089976606775093909

Rapport, M. D., Orban, S. A., Kofler, M. J., and Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clin. Psychol. Rev* 33, 1237–1252. doi: 10.1016/j.cpr.2013.08.005

Rebok, G. W., Carlson, M. C., and Langbaum, J. B. (2007). Training and maintaining memory abilities in healthy older adults: traditional and novel approaches. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 62, 53–61. doi: 10.1093/geronb/62.special_issue_1.53

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *J. Exp. Psychol. Gen.* 142, 359–379. doi: 10.1037/a0029082

Romine, C. B., and Reynolds, C. R. (2005). A model of the development of frontal lobe function: findings from a meta-analysis. *Appl. Neuropsychol.* 12, 190–201. doi: 10.1207/s15324826an1204_2

Rose, S. A., Feldman, J. F., and Jankowski, J. J. (2011). Modeling a cascade of effects: the role of speed and executive functioning in preterm/full-term differences in academic achievement. *Dev. Sci.* 14, 1161–1175. doi: 10.1111/j.1467-7687.2011.01068.x

Rossi, A. F., Pessoa, L., Desimone, R., and Ungerleider, L. G. (2009). The prefrontal cortex and the executive control of attention. *Exp. Brain Res.* 192, 489–497. doi: 10.1007/s00221-008-1642-z

Rubia, K., Russell, T., Overmeyer, S., Brammer, M. J., Bullmore, E. T., Sharma, T., et al. (2001). Mapping motor inhibition: conjunctive brain activations across different versions of go/no-go and stop tasks. *Neuroimage* 13, 250–261. doi: 10.1006/nimg.2000.0685

Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E., et al. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Hum. Brain Mapp.* 27, 973–993. doi: 10.1002/hbm.20237

Rueda, M. R., Rothbart, M. K., McCandliss, M. D., Saccomanno, L., and Posner, M. L. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14931–14936. doi: 10.1073/pnas.0506897102

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology* 19, 532. doi: 10.1037/0894-4105.19.4.532

Scherf, K. S., Sweeney, J. A., and Luna, B. (2006). Brain basis of developmental change in visuospatial working memory. *J. Cogn. Neurosci.* 18, 1045–1058. doi: 10.1162/jocn.2006.18.7.1045

Schuchardt, K., Maehler, C., and Hasselhorn, M. (2008). Working memory deficits in children with specific learning disorders. *J. Learn. Disabil.* 41, 514–523. doi: 10.1177/0022219408317856

Shah, P., Buschkuehl, M., Jaeggi, S., and Jonides, J. (2012). Cognitive training for ADHD: the importance of individual differences. *J. Appl. Res. Mem. Cogn.* 1, 204–205. doi: 10.1016/j.jarmac.2012.07.001

Shallice, T. I. M., and Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain* 114, 727–741. doi: 10.1093/brain/114.2.727

Shipstead, Z., Redick, T. S., and Engle, R. W. (2012). Is working memory training effective? *Psychol. Bull.* 138, 628–654. doi: 10.1037/a0027473

Sowell, E. R., Thompson, P. M., Tessner, K. D., and Toga, A. W. (2001). Mapping continued brain growth and gray matter density reduction in dorsal frontal cortex: inverse relationships during postadolescent brain maturation. *J. Neurosci.* 21, 8819–8829.

Strobach, T., Salminen, T., Karbach, J., and Schubert, T. (2014). Executive functions and their optimization in the context of dual-task practice. *Psychol. Res.* doi: 10.1007/s00426-014-0563-7

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Stuss, D. T., and Benson, D. F. (1986). *The Frontal Lobes.* New York: Raven Press.

Swanson, H. (2004). Working memory and phonological processing as predictors of children's mathematical problem solving at different ages. *Mem. Cogn.* 32, 648–661. doi: 10.3758/BF03195856

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., and Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12, 106–113. doi: 10.1111/j.1467-7687.2008.00745.x

Titz, C., and Karbach, J. (2014). Working memory and executive functions: effects of training on academic achievement. *Psychol. Res.* doi: 10.1007/s00426-013-05371 [Epub ahead of print].

Unsworth, N., and Engle, R. W. (2006). A temporal-contextual retrieval account of complex-span: an analysis of errors. *J. Mem. Lang.* 54, 346–362. doi: 10.1016/j.jml.2005.11.004

van Merriënboer, J. G. J., Kester, L., and Paas, F. (2006). Teaching complex rather than simple tasks: balancing intrinsic and germane load to enhance transfer of learning. *Appl. Cogn. Psychol.* 20, 343–352. doi: 10.1002/acp.1250

Verhaeghen, P. (2014). *The Elements of Cognitive Aging: Meta-Analyses of Age-Related Differences in Processing Speed and Their Consequences.* Oxford: Oxford University Press.

Verhaeghen, P., Marcoen, A., and Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: a meta-analytic study. *Psychol. Aging* 7, 242. doi: 10.1037/0882-7974.7.2.242

von Bastian, C. C., and Oberauer, K. (2013). Distinct transfer effects of training different facets of working memory capacity. *J. Mem. Lang.* 69, 36–58. doi: 10.1016/j.jml.2013.02.002

Wager, T. D., Jonides, J., and Reading, S. (2004). Neuroimaging studies of shifting attention: a meta-analysis. *Neuroimage* 22, 1679–1693. doi: 10.1016/j.neuroimage.2004.03.052

Wager, T. D., and Smith, E. E. (2003). Neuroimaging studies of working memory. *Cogn. Affect. Behav. Neurosci.* 3, 255–274. doi: 10.3758/CABN.3.4.255

Wendelken, C., Bunge, S. A., and Carter, C. S. (2008). Maintaining structured information: an investigation into functions of parietal and lateral prefrontal cortices. *Neuropsychologia* 46, 665–678. doi: 10.1016/j.neuropsychologia.2007.09.015

Wendelken, C., Munakata, Y., Baym, C., Souza, M., and Bunge, S. A. (2012). Flexible rule use: common neural substrates in children and adults. *Dev. Cogn. Neurosci.* 2, 329–339. doi: 10.1016/j.dcn.2012.02.001

Wiebe, S. A., Espy, K. A., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575. doi: 10.1037/0012-1649.44.2.575

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., and Espy, K. A. (2011). The structure of executive function in 3-year-olds. *J. Exp. Child Psychol.* 108, 436–452. doi: 10.1016/j.jecp.2010.08.008

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends Cogn. Sci.* 8, 12–17. doi: 10.1016/j.tics.2003.11.001

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Zelazo, P. D., and Lyons, K. E. (2012). The potential benefits of mindfulness training in early childhood: a developmental social cognitive neuroscience perspective. *Child Dev. Perspect.* 6, 154–160. doi: 10.1111/j.1750-8606.2012.00241.x

Zinke, K., Einert, M., Pfennig, L., and Kliegel, M. (2012). Plasticity of executive control through task switching training in adolescents. *Front. Hum. Neurosci.* 6:41. doi: 10.3389/fnhum.2012.00041

Zinke, K., Zeintl, M., Rose, N. S., Putzmann, J., Pydde, A., and Kliegel, M. (2014). Working memory training and transfer in older adults: effects of age, baseline performance, and training gains. *Dev. Psychol.* 50, 304–315. doi: 10.1037/a0032982

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 February 2014; accepted: 14 April 2014; published online: 07 May 2014.
*Citation: Karbach J and Unger K (2014) Executive control training from middle childhood to adolescence. Front. Psychol. **5**:390. doi: 10.3389/fpsyg.2014.00390*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Karbach and Unger. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 13 March 2014
doi: 10.3389/fpsyg.2014.00205

Predictors of cognitive enhancement after training in preschoolers from diverse socioeconomic backgrounds

M. Soledad Segretin[1]*, Sebastián J. Lipina[1], M. Julia Hermida[1], Tiffany D. Sheffield[2,3], Jennifer M. Nelson[2,3], Kimberly A. Espy[3,4] and Jorge A. Colombo[1]

[1] Unidad de Neurobiología Aplicada, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Centro de Educación Médica e Investigaciones Clínicas Norberto Quirno (CEMIC), Ciudad Autonoma de Buenos Aires, Buenos Aires, Argentina
[2] Office of Research, University of Nebraska, Lincoln, NE, USA
[3] Department of Psychology, University of Nebraska, Lincoln, NE, USA
[4] Department of Psychology, University of Oregon, Eugene, OR, USA

Edited by:
Philip D. Zelazo, University of Minnesota, USA

Reviewed by:
Silvia A. Bunge, University of California Berkeley, USA
Frederick James Morrison, University of Michigan, USA

***Correspondence:**
M. Soledad Segretin, Unidad de Neurobiología Aplicada, Consejo Nacional de Investigaciones Científicas y Técnicas, Centro de Educación Médica e Investigaciones Clínicas, Unidad de Neurobiología Aplicada, Avenida Galván 4102, Ciudad Autónoma de Buenos Aires C1431FWO, Argentina
e-mail: soledadsegretin@gmail.com

The association between socioeconomic status and child cognitive development, and the positive impact of interventions aimed at optimizing cognitive performance, are well-documented. However, few studies have examined how specific socio-environmental factors may moderate the impact of cognitive interventions among poor children. In the present study, we examined how such factors predicted cognitive trajectories during the preschool years, in two samples of children from Argentina, who participated in two cognitive training programs (CTPs) between the years 2002 and 2005: the *School Intervention Program* (SIP; $N = 745$) and the *Cognitive Training Program* (CTP; $N = 333$). In both programs children were trained weekly for 16 weeks and tested before and after the intervention using a battery of tasks assessing several cognitive control processes (attention, inhibitory control, working memory, flexibility and planning). After applying mixed model analyses, we identified sets of socio-environmental predictors that were associated with higher levels of pre-intervention cognitive control performance and with increased improvement in cognitive control from pre- to post-intervention. Child age, housing conditions, social resources, parental occupation and family composition were associated with performance in specific cognitive domains at baseline. Housing conditions, social resources, parental occupation, family composition, maternal physical health, age, group (intervention/control) and the number of training sessions were related to improvements in specific cognitive skills from pre- to post-training.

Keywords: cognitive development, intervention, SES, mixed models, socio-environmental predictors, preschool children

INTRODUCTION

Broadly defined, executive functions (EF) refer to a complex set of cognitive abilities that underlie adaptive, goal-directed behaviors, and enable individuals to override more automatic or established thoughts and responses (Garon et al., 2008; Diamond, 2013). EF are critical when solving novel problems and thus essential for self-regulation, school learning, and social behavior (e.g., Hughes and Graham, 2002; Anderson, 2002; Isquith et al., 2005; Diamond et al., 2007; Garon et al., 2008; Bull et al., 2011; Espy et al., 2011). At a more fine-grained level a set of cognitive control skills (e.g., attention, inhibitory control, self-monitoring, and flexibility) is defined as specific interrelated information-processing abilities that are involved in the control and coordination of information in the service of goal-directed actions, as studied in the cognitive development literature (Willoughby et al., 2012). Focusing on these more narrowly defined abilities is particularly suitable when studying EF in early childhood, as many of the more complex aspects of EF (e.g., abstract thought; goal setting) have an extended developmental course and are not easily measured in very young children (Garon et al., 2008; Willoughby et al., 2012). The emergence and development of those cognitive processes depend on both biological maturation and environmental experiences (Fisher, 2006; Berkman et al., 2012), and follow different trajectories from the first year of life (Anderson, 2002; Garon et al., 2008; Marsh et al., 2010; Howe et al., 2012). In addition, these trajectories are sensitive to individual differences, and the quality of the micro- and mesosystemic developmental contexts (home and school) (Lipina and Colombo, 2009; Cadima et al., 2010; Sarsour et al., 2011).

Several studies have suggested that associations between socioeconomic status and cognitive development during childhood are mediated by biological, psychological and environmental factors, which may be conceptualized at multiple levels of analysis (individual, family, and social contexts), and increase the likelihood of negative impacts later in life (Leinonen et al., 2002; Raver et al., 2007, 2013; Santos et al., 2008; Cadima et al., 2010; Rhoades et al., 2011; Sarsour et al., 2011; Lipina et al., 2013). Among the environmental factors that have been associated with these impacts, the following are the most cited in the scientific literature: family income, family composition, parental level of education and occupation, housing conditions, perinatal health factors, quality of home and school environments, attendance

to early education programs, parental mental health, parenting styles, parent-child interactions, neighborhood characteristics, and social support (Burchinal et al., 2000; Bradley and Corwyn, 2002, 2005; Evans, 2004; Gassman-Pines and Yoshikawa, 2006; Engle et al., 2007; Grantham-McGregor et al., 2007; Walker et al., 2007; Rhoades et al., 2011; Sarsour et al., 2011). Additionally, the impact of these factors on different aspects of child development may vary according to the type, number and accumulation of risk factors to which children are exposed, the timing of exposure, and the individual susceptibility to each one (Najman et al., 2004; Stanton-Chapman et al., 2004; Belsky et al., 2007; Walker et al., 2007; Kiernan and Huerta, 2008; Flouri et al., 2009; Kiernan and Mensah, 2009; Hall et al., 2010; Rhoades et al., 2011; Evans et al., 2013). Thus, it is not only the mere presence or absence of specific risk factors that influence development, but also their accumulation in a context of individuality, with more risk leading to greater adjustment difficulties (Burchinal et al., 2000; Stanton-Chapman et al., 2004; Appleyard et al., 2005; Gassman-Pines and Yoshikawa, 2006; Cadima et al., 2010; Evans et al., 2013). Despite the significant advances in the field, more research is necessary to elucidate specific environmental experiences that contribute to individual differences in cognitive control development (Rhoades et al., 2011), as well as their contribution to individual differences in the context of intervention trials.

Specifically, cognitive control performance of children living in poverty is limited in its potential due to the presence of multiple risk factors in these contexts, such as child health history (peri- and postnatal), maternal education, parental mental health, quality of stimulation at home, and social interactions in different contexts (e.g., home and school). Results from studies developed in Argentina assessing associations between poverty and impact on cognitive processing have verified the modulation of different cognitive processes (i.e., attentional, inhibitory control, working memory, flexibility, and planning) in infants and preschoolers as a result of socioeconomic status and income, as well as the influence of poverty on academic performance (i.e., language and mathematics) in elementary and high school children (Lipina et al., 2004, 2005, 2013; Segretin et al., 2009).

In Argentina, according to the latest data published by the National Institute of Statistics and Censuses (INDEC, 2001), 14.3% of families live with Unsatisfied Basic Needs (UBN). Additionally, the Observatory on the Argentinean Social Debt reported that 29.6% of the population was poor during 2010 (Tuñón, 2011). With respect to child poverty, 40.5% of children under the age of 14 were living in poverty, and 14.2% in extreme poverty in 2006 (INDEC, 2006). According to the report of the United Nations Fund for Children of 2010, 28.7% of children under the age of 18 in Argentina are in poverty (CEPAL-UNICEF, 2010).

During the past decade, several interventions targeting cognitive control development have been designed and evaluated in the fields of developmental psychology and developmental cognitive neuroscience (Lipina and Colombo, 2009; Burger, 2010). The main goals of such interventions were the promotion of cognitive control development in early childhood, with the aim of influencing broader, long-term outcomes, such as academic and social adjustment (McCandliss et al., 2003; Temple et al., 2003; Colombo and Lipina, 2005; Klingberg et al., 2005; Rueda et al., 2005; Wilson et al., 2006; Diamond et al., 2007; Stevens et al., 2008; Beatty, 2009; Thorell et al., 2009; Barnett, 2011; Espinet et al., 2012). Most of the previous interventions were successful in promoting cognitive performance in the short- or medium-term, and evaluation of their success generally focused on pre- and post-training performance comparisons between groups. Only few studies have also included an analysis of the predictors of intervention impact, mostly based on variables, such as the initial cognitive performance, age, and/or program characteristics (e.g., Bierman et al., 2008). The effectiveness of the intervention programs that include cognitive stimulation modules has been related to the following aspects of program design: (a) comprehensiveness of services (educational, nutritional, sanitation, and social services); (b) teacher and family participation; (c) direct and indirect interventions; (d) quality of services; (e) staff recruitment and training; and (f) cultural pertinence of interventions (Ramey and Ramey, 1998, 2003; Gray and McCormick, 2005; Karoly et al., 2005; Reynolds and Temple, 2005; Perez-Johnson and Maynard, 2007; Barnett, 2011; Reynolds et al., 2011). It is important, however, to consider that interventions are not equally effective for all participants. In this regard, different factors could moderate the impact of the intervention, and this information would be crucial to the design of new interventions, both experimental and applied.

In general, studies on the moderation of cognitive development by environmental factors have focused on the associations between child poverty and accumulation of risk factors (Gassman-Pines and Yoshikawa, 2006; Weiland and Yoshikawa, 2012; Flouri et al., 2013; Meunier et al., 2013). Less analytical efforts have been devoted to moderation based on risk factors in the area of intervention science. Thus, the implementation of risk-factor analysis, such as identifying different socio-environmental variables as predictors of intervention impact, is important in order to establish targets for improvement in the design of innovative interventions.

In this article, we propose the analysis of specific aspects of two intervention programs implemented in Argentina between 2002 and 2005, with the main goal of optimizing cognitive control performance in preschool children: the *School Intervention Program* (SIP), and the *Cognitive Training Program* (CTP) (**Table 1**). We focus the analysis on the identification of different socio-environmental predictors of cognitive trajectories. Specifically, the main goals of the present study were: (1) to examine how environmental factors moderate cognitive performance; and (2) to identify factors that moderate the impact of two intervention programs aimed at optimizing cognitive performance in two samples of poor- and non-poor preschoolers. We examined children's performance in tasks demanding attention, inhibitory control, memory, flexibility, and planning, and considered the impact of environmental risk factors on each cognitive task at baseline and task trajectories (change in performance from pre- to post-assessment). Mixed model analyses were applied in order to identify socio-environmental predictors associated with higher levels of cognitive control performance in the pre-intervention phase, and with increased improvement in cognitive control between pre- and post-intervention phases. We expected that higher cognitive performance at baseline would be associated with better

Table 1 | Programs descriptions.

	Program	
	SIP	**CTP**
Design	Experimental, controlled, and random study.	Quasi-experimental and random study.
Participants	Children from 3 to 5 years from UBN[a] homes. City of Buenos Aires.	Children from 3 to 5 years from UBN and SBN[b] homes. City of Salta.
Study groups	Intervention/control.	Individual training modality/group training modality*.
Program phases	1. Cognitive assessment (Time 1) 2. Intervention modules implementation 3. Cognitive assessment (Time 2)	1. Cognitive assessment (Time 1) 2. Intervention modules implementation 3. Cognitive assessment (Time 2)
Intervention modules	Cognitive training; nutritional supplementation; counseling for parents; training; and counseling for teachers.	Cognitive training; nutritional supplementation (the government agency provided counseling for parents and adults working in the childcare centers).
Activities for the cognitive training module	Exercising activities (individual modality of training).	Activities with a pedagogical format (individual and group modalities of training).
Frequency of intervention	Once a week for 16 weeks in 1 year; twice a week during 16 weeks in 1 year.	Twice a week for 16 weeks in 1 year.
Total cognitive training sessions	16 or 32 sessions.	32 sessions.
Context of implementation	Kindergartens.	Childcare Centers.

Only 4-year-old children were randomly assigned to individual or group training modalities. For that reason, analysis were run separately for this age group (with the aim to compare training modalities) (manuscript under revision), and in the present article only 3- to 5-year-old children assigned to the group training modality were considered for the prediction analysis.

[a] UBN (poverty criteria, see details in section Socioeconomic, Life, and Health Condition Measures). [b] SBN: Satisfied Basic Needs.

socio-environmental conditions (Feldman and Eidelman, 2009; Kiernan and Mensah, 2009). Additionally, we expected that children living in families with more resources (in terms of parental occupation and education, financial resources, type of housing and social support) would have higher improvements in their cognitive performance after training. This hypothesis was based on the idea that for children from worst socio-environmental conditions another type of intervention would be required (e.g., more specific for each cognitive skill, with more intervention intensity exposure-considering frequency, length, and age) (Ramey and Ramey, 2003). We also expected to identify different predictors for each cognitive process and program, taking into account the differences in the cognitive developmental trajectories at these ages (Garon et al., 2008), and the differences in the program characteristics (e.g., context of implementation, number and frequency of training sessions, and modalities of training) (Jolles and Crone, 2012). Also, it is important to mention that results from different intervention programs for disadvantaged children have suggested that the frequency of intervention is a significant modulator of the impact (Ramey and Ramey, 2003; Karoly et al., 2005; Burger, 2010). Based on this, we also expected that the number of training sessions (exposure to training) would be a significant predictor of cognitive trajectories, with higher cognitive performance improvements in children with more exposure to training.

METHODS

STUDY DESIGN, PARTICIPANTS, AND PROCEDURES

SIP program

A sample of healthy Argentinean children aged 3–5 years participated in the SIP, a longitudinal study implemented between years 2002 and 2004 in three kindergartens in the city of Buenos Aires, selected by applying the conglomerate sample method. The program was an experimental, randomized and controlled study with the main goal to train cognitive performance in preschool children from UBN homes (Colombo and Lipina, 2005; Lipina et al., 2012). Seven hundred and forty five preschool children were authorized to participate. We have verified an attrition rate around 15% per year. Each year new cohorts were enrolled, forming different study groups with 1 and 2 years of intervention. For the present article only data from children with 1 year of intervention between the years 2002 and 2004 were analyzed, because of the small sample size for 2 years of intervention, which did not allow executing the planned analytical procedures. Informed consents were obtained from parents/caregivers, and ethical approval was obtained from the Ethical Review Committee. The study was conducted according to APA's ethical standards, and international and national children rights laws.

Before the beginning of the program, we recruited and trained a group of college students ("trainers" from now on) from the school of psychology and education. During the same period, we informed the parents of children attending the selected institutions about the program activities and we asked them to sign written consents to include their children into the program. After that, from April to July in each year, we evaluated children's cognitive performance (Time 1, baseline) (see section Cognitive Measures), and parents attended individual interviews to give socioeconomic, sociodemographic and child health history information. Then, four intervention modules were implemented from July to November (see above). After the intervention, all children had a final cognitive assessment (Time 2, post-intervention) with the same battery of tasks used at Time 1. We provided trainers with an intervention procedure guide and supervised them daily. Trainers had to complete a form describing the implemented activities, which we revised daily to suggest adjustments. Trainers were blind to the hypotheses of the study. Activities for the control and intervention groups were organized in different days of the week

and all trainers were reassigned to different schools for the final phase (post-intervention cognitive assessment).

The program included the following four modules:

(1) The individual *cognitive training module* consisted of different exercises demanding cognitive control, with increasing complexity. Activities involved some of the tasks used for the pre- and post-assessment, using different trials (**Figure 1**). Exercises were implemented during the school day as an extracurricular activity. Two schemes were applied: once a week (a total of 16 sessions), or twice a week (a total of 32 sessions) during 16 weeks in 1 year. Activities followed a scheme previously designed considering the cognitive demand and the time available for the session (30/40 min), and were implemented by one trainer (adults/children ratio = 1 per 1), who was the same for each child.

Each training session was structured in four consecutive steps: (1) assessment of children's motivational state with a Likert scale including the following constructs: willing/not willing to collaborate, extroversion/introversion, talkative/quiet, active/passive, impulsive/thoughtful, trustful/distrustful; (2) introduction of novel materials for the activity and task instructions; (3) evaluation of instruction comprehension with pretest exercises; and (4) activities: blocks of different exercises or trials. Only when children were adequately motivated and pretests were properly solved, trainers continued with the next step, otherwise, the activity was scheduled for a new day in the same week. With respect to the fourth step, for each activity children were asked to solve different exercises organized in blocks of 5–10 trials (two blocks of exercises per session). After each block of exercises was completed, trainers evaluated children's performance, and determined the complexity level for the second block. When child efficiency in the first block of exercises reached at least 80%, the trainers increased the complexity level for the next block; otherwise, after trainers' intervention (according to the child's difficulties, considering the problem solving scheme), the second block of exercises presented new trials with the same level of difficulty. Performance on the last block of exercises determined the initial level of difficulty in the next session (**Figure 2**).

The theoretical framework of the cognitive training module was based on the problem-solving framework proposed by Zelazo et al. (1997). It involves four temporally and functionally distinct steps and substeps: problem representation, planning, execution, and evaluation (detection and correction). To solve a problem, first it is necessary to create or restructure the problem representation, including its possible solutions. Another component considered in the cognitive training module was the dynamic testing approach proposed by Sternberg and Grigorenko (2002). This form of testing proposes giving the children some kind of feedback in order to help them improve their scores, which in turn is based on Vigotsky's conception of the proximal development zone. Finally, two other components of the cognitive training module were the inclusion of challenge activities or trials, and repeated practice (Diamond, 2012).

(2) The *nutritional supplementation module* (implemented for both groups) consisted of the administration of one pill per week during the cognitive training period. Each pill contained 60 mg of elementary iron and 0.4 mg of folic acid, and was provided by UNICEF-Argentina.

(3) *Parental Counseling* was a module implemented for both groups, throughout the school year. Activities in this module included: (a) parent counseling; (b) child clinical exam; (c) child blood extraction to identify levels of hemoglobin; and (d) parent interviews.

(4) *Teacher training and counseling* (for both groups) was the fourth module, implemented throughout the school year, twice a month.

CTP Program

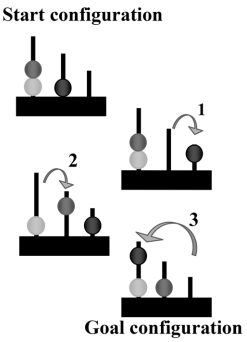
Based on the results of the *SIP*, the same group of researchers designed a new CTP. In the year 2005, this program was implemented in the city of Salta in the context of a quasi-experimental prospective design (Segretin et al., 2007a,b; Lipina et al., 2012). Specifically, the aims of the program involved fostering cognitive development in preschoolers from UBN and Satisfied Basic Needs (SBN) homes with a reduction in the adult to child ratio (more children per adult) compared to the previous experience (SIP) (1/15 vs. 1/1, respectively). For this longitudinal study, a sample of 382 healthy Argentinean children aged 3–5 years were recruited from official childcare centers in the city of Salta in Argentina (Secretary of Children and Family from the Government of the Province of Salta) applying a conglomerate sample method. The rate of attrition was 15%. Informed consents were obtained from parents/caregivers, and ethical approval was obtained from the Ethical Review Committee. The study was conducted according to APA's ethical standards, and international and national children rights laws.

In this program, we recruited and trained a group of trainers, and we informed parents of children attending the selected institutions about the program activities and asked them to sign written consents to include their children into the program. After that, from April to July, we evaluated children's cognitive performance (Time 1) (see section Cognitive Measures), and parents attended individual interviews to give socioeconomic, sociodemographic and child health history information. Then, 4-year-old children were randomly assigned to an individual or group modality of cognitive training. Three- and 5-year-old children were all assigned to the group modality. The reasons for such design were: (1) authorities did not allow the research team to generate a control group for ethical reasons (i.e., they considered that all children had to receive the same activities, and that the government was not a research agency aimed at supporting research practices); and (2) authorities required to reduce the number of human resources for the execution of the program (i.e., individual training modality requires more trainers). For the present study, only children assigned to the group modality of training were analyzed ($n = 333$). The rest of the children (49) were trained with the individual modality, which was similar to the one implemented in the SIP (Colombo and Lipina, 2005; Lipina and Segretin, 2006; Martelli et al., 2007), to have a comparative training group (these children are not considered further in this paper) (Segretin et al., submitted for

Training activities (SIP)

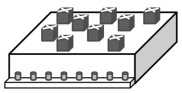
Planning training

Tower of London: In each trial of this task, the child was required to reach a goal configuration of three colored balls from a start configuration, moving one ball at a time, in a minimum number of moves. Levels of difficulty comprised exercises from one to nine moves, and each level comprised 5 trials. Any colored ball might be placed on top of any other colored ball, and children had to generate the appropriate action sequence. Trials were administered until the child made 3 consecutive errors. For example, three movements are required to reach the goal configuration in this example:

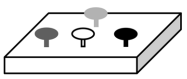


Working Memory training

Corsi Blocks: In this task, a sequence of lights was turned on (two seconds each light, one second between lights). The child was required to remember and point the boxes following the light sequences. Levels of difficulty comprised exercises from two to eight lights, and each level comprised 5 trials. The task was administered until the child made 3 consecutive errors.



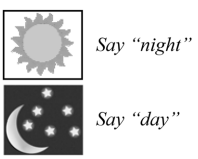
Three and four colors: It consisted of a box with four colored buttons. After the experimenter told the child a sequence of colors (between 2 and 6), the child was required to reproduce the sequence, pressing the same colored buttons. Each block of exercises consisted of 2-to-7 trials, according to the child´s performance.



Digit Span: In this task, the child was required to listen, remember and reproduce a sequence of numbers (between 2 and 9). Each block consisted of 5 trials with the same length of span.

Inhibitory control training

Stroop-like Day-Night: In this task, children were asked to say the opposite of what they saw in a series of cards (for example, when a picture of a sun was presented, they had to say "night", and when the picture showed was a moon, they had to say "day"). For each complexity level, a set of 16 cards were presented to the child. In the easier level, only two different cards were used (one opposite pair), whereas in the more difficult level, six different cards were mixed (three opposite pairs).



Categorization and flexibility training

Flexible Item Selection Task (FIST): A set of 36 cards were used in this task, conformed with items according to four dimensions: shape, color, size, and number. Task consisted of 12 trials in which the child was required to select two cards sharing the same criteria. In the easier level, the child was required to make one selection (selecting two cards), but in the second level, the child was required to make two selections: first, select two cards equals in one way, and then select another other two cards in a different way. In the example the first and the third cards are equals because they have two elements. Likewise, the first and the second cards are equals because they have fishes. Size and color are equals in all cards; number and shape are common in two cards.

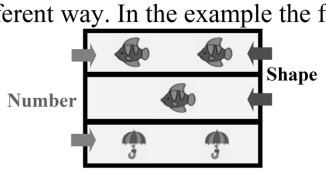


FIGURE 1 | Description of training activities in the SIP.

publication). Then, from July to November two intervention modules were implemented (see above), and after that, all children were administered a final cognitive assessment (Time 2) with the same battery of tasks used at Time 1. Like in the previous program, trainers' work in each phase was supervised daily during the year, and they were provided a procedure guide. Also, trainers had to complete a form for each activity, which were reviewed daily by supervisors.

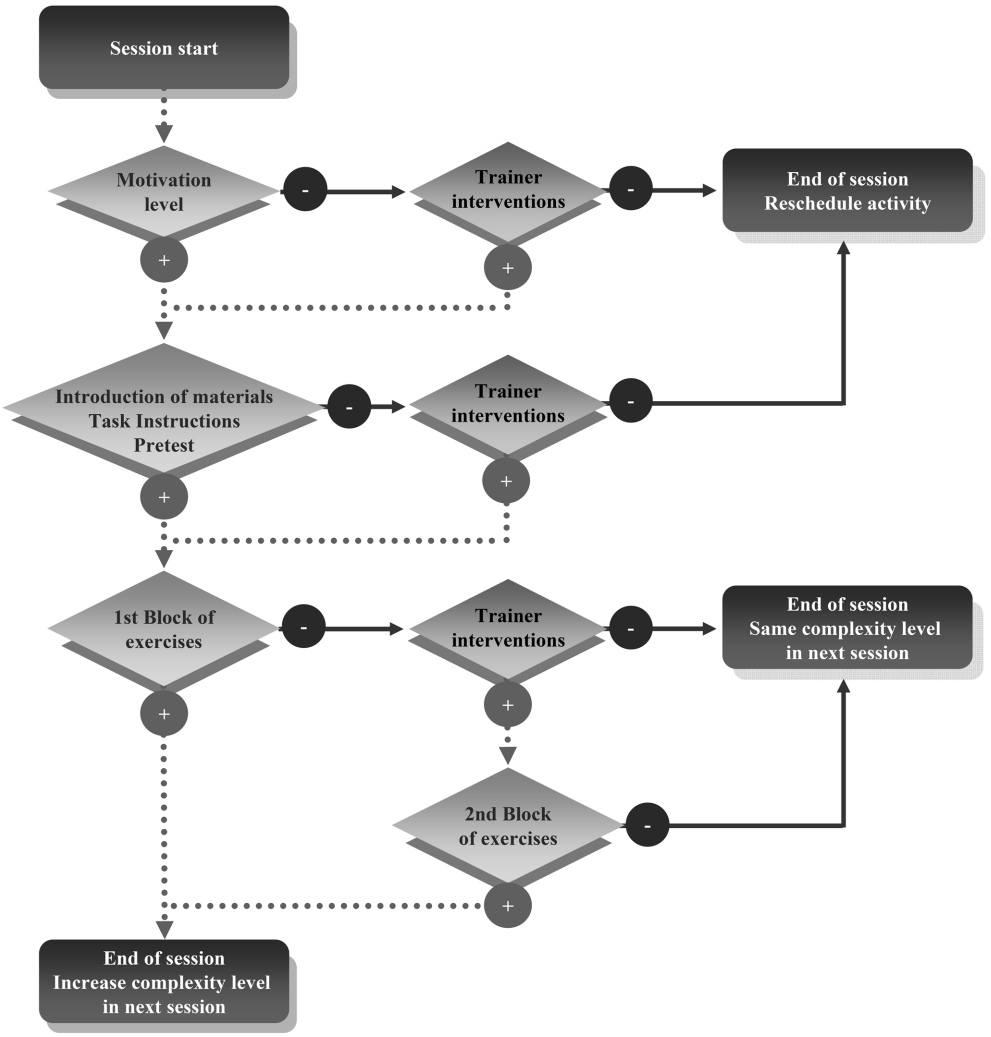


FIGURE 2 | Flowchart of the structure of each training session.

In the CTP two modules were implemented, and put together with other activities developed by the government agency:

(1) The *cognitive training module* consisted of different activities demanding cognitive control, with increasing complexity. Activities for both training modalities were designed with a group of pedagogues (in this program activities differed from the basal and post-training cognitive assessment), and were implemented in weekly sessions (with two different activities within each session) during the school day, as an extracurricular classroom activity, for 16 weeks. As previously mentioned, for the present work only children in the group modality of training were considered in the analysis. Training groups were organized based on age and the maximum number of children necessary to form each group (from 10 to 25 children). Activities followed a scheme previously designed considering the cognitive demands and the time available for each session (30/40 min), and were implemented by two trainers (adults/children ratio = 1 per 10/15) (**Figure 3**).

The theoretical framework for the cognitive training module was the same one that was used in the SIP program. Also, the CTP applied the same session structure as the previous program (see **Figure 2**), but adapted to the group modality of training. That is, activities were solved with the participation of all children in the group, and only when 80% of children successfully solved 80% of trials in one block of exercises was the complexity level increased.

(2) The nutritional supplementation module was implemented for all children in the program, with the same frequency and methodology applied in the SIP.

Additionally, the government agency provided counseling for parents and adults working in the childcare centers. Researchers had no access to the information regarding these interventions.

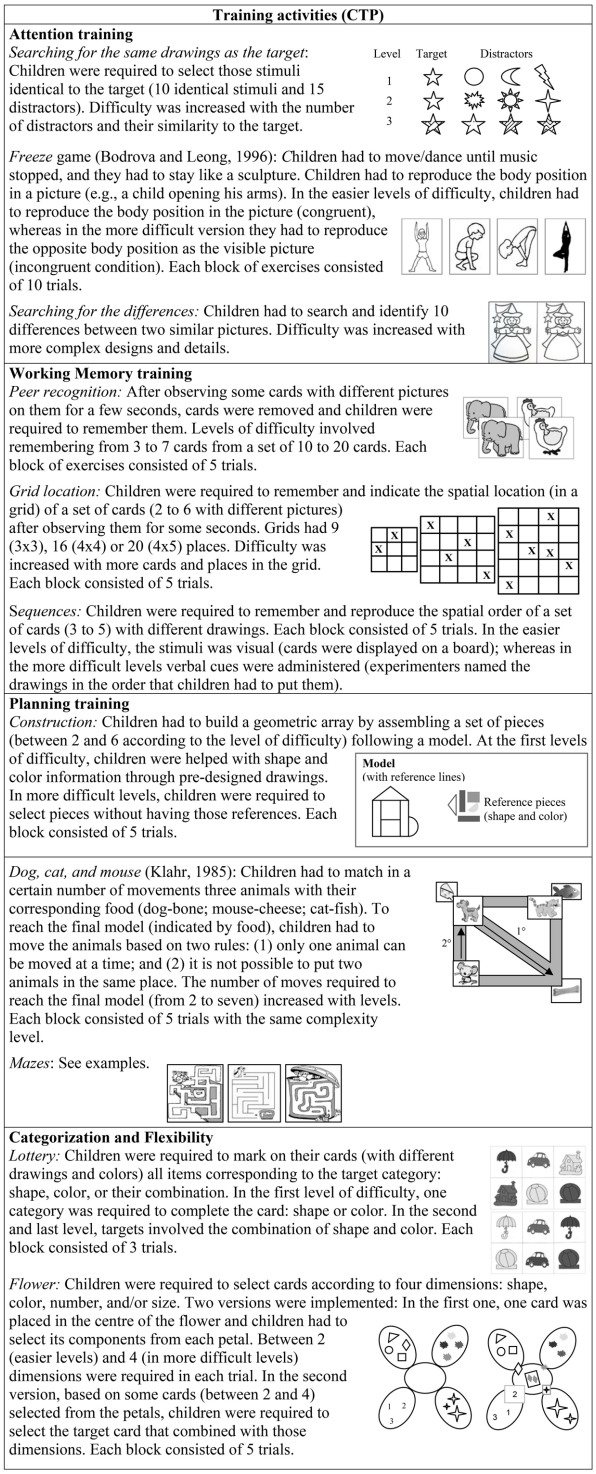


FIGURE 3 | Description of training activities in the CTP.

MEASURES IN BOTH PROGRAMS

Socioeconomic, life, and health condition measures

In both programs data were collected during the school year (March to November) in a private interview with parents. A Socioeconomic Scale (NES) was used to evaluate parents' education and occupation levels, overcrowding, housing and sanitation conditions, to identify indicators of UBN (Boltvinik, 1995). Scores were assigned directly to mothers and fathers for educational and occupational backgrounds; however, only the higher score was considered for the total scores. For housing conditions, scores were assigned based on type of dwelling, floor, water, bathroom, ceiling, external walls, and home property. The Life Stressors and Social Resources Inventory (LISRES inventory) (Mikulic, 1999) was used to identify life stressors and social resources in the family. Specifically, this inventory measures sets of stressors and resources by administering two scales: (1) *stressors* that includes physical health dimensions (29 items), housing/neighborhood (22 items), finance (9 items), work (15 items), family (13 items), children (15 items), extended family (15 items), friends and social activities, and negative life events (12 items); and (2) *resources* that includes: finance (9 items), work (15 items), family (13 items), children (15 items), extended family (15 items), friends and social activities (15 items), and positive life events (10 items). The total score for each scale was calculated by adding scores obtained in each set of items, which were then transformed into *T*-scores (mean = 50, standard deviation = 10). Additionally, a set of questions concerning child peri- and post-natal health conditions was included in the interviews. Finally, in the SIP the Hamilton Scale (Hamilton, 1959, 1960) was employed to consider important aspects of mothers' mental health involved in self-regulation development (Buss et al., 2011). The scale consists of 14 items related to signs and symptoms of anxiety (7 items) and depression (7 items), which measures the intensity and frequency of such behaviors. The sum of the specific items for each type of sign results in a total score for depression and another for anxiety. There are no cut-points to distinguish subjects with and without anxiety or depression, so the results should be interpreted as a quantification of the intensity.

Cognitive measures

We evaluated cognitive performance in the pre- and post-training phases (Time 1 and Time 2), with a set of tasks administered by trainers in two sessions of about 40 min. Children were tested individually at their schools, in a quiet testing room. Testing was scheduled at times reported by teachers not to interfere with regular meals and activities. Examiners were blind to the objectives of the study and the composition of the groups.

In both programs, the following four tasks were used:

(a) The *Selective Attention* task—a manual adaptation of the computerized version of a subscale of the NEPSY battery (Korkman et al., 1998)—was used to evaluate attentional control. A set of sheets of paper with 25 pictures and one or more targets on each one was used. The child was required to identify and point to all the drawings that were identical to the target. Levels of difficulty (from 1 to 10) were determined according to the number of targets and the similarity between the target and the distractor drawings. Trials were administered until the child made more than 3 errors and/or omissions in three consecutive sheets. Scores represent the proportion of correct responses.

(b) The *Corsi Blocks* task (Berch et al., 1998) was used to assess visuo-spatial organization processes. In this task, a sequence of lights (from 2-to-8) was turned on (2 s each light, 1 s between lights). The child was required to remember and point to the boxes following the light sequences. Each level of difficulty comprised five trials with the same number of elements that children had to remember. Trials were administered until the child made three consecutive errors. A total score was computed as the sum of correct responses multiplied by level of difficulty (determined by the number of elements to remember on each trial).

(c) The *Tower of London* task (Shallice, 1982) was used to assess planning. In each trial the child was required to reach a goal configuration of three colored balls from a start configuration, moving one ball at a time, in a minimum number of movements. Any colored ball might be placed on top of any other ball, and children had to generate the appropriate action sequence. Levels of difficulty comprised exercises from 1 to 9 movements, and each one comprised five trials. Trials were administered until the child made three consecutive errors. A total score was computed as the sum of correct responses multiplied by level of difficulty (determined by the minimum number of movements necessary to reach the final model).

(d) The *Flexible Item Selection Task* (Jacques and Zelazo, 2001) was administered in the CPT, and in the second and third year of the SIP—which explains the reduction of the sample size—to assess abstract processing and cognitive flexibility. A set of cards was used conformed to items according to four dimensions: shape, color, size, and number. Tasks consisted of 12 trials in which the child was required to select two cards sharing the same dimension. The child was required to make two selections: first, select two cards equivalent in one way, and then select another two cards, equivalent in a different way. A total score was computed by adding correct responses in the first and second selection.

In addition, in the SIP, the *Stroop-like Day-Night* task was administered to assess inhibitory control processes (Gerstadt et al., 1994). The task consisted of 16 trials in which children were asked to say the opposite of what they saw in a series of cards. When a picture of a *sun* was presented, they had to say "night," and when the picture showed a *moon*, they had to say "day." A total score was computed as the sum of correct responses divided by the total number of trials.

RESULTS

SIP PROGRAM

Based on the literature, a set of variables were pre-selected as potential predictors of cognitive performance at baseline and of the change in performance between pre- and post-intervention: *housing conditions, overcrowding, parental education, parental occupation, mother's physical health, housing stressors, economic stressors, working stressors, couple stressors, child stressors, family stressors, friends, and social life stressors, negative life events, economic resources, working resources, couple resources, child resources, family resources, friends, and social life resources, positive life events, child health records, child age, child gender*, and *frequency of training sessions* (Sameroff et al., 1993; Brooks-Gunn and Duncan, 1997; McLoyd, 1998; Burchinal et al., 2000; Bradley and Corwyn, 2002; Gassman-Pines and Yoshikawa, 2006). Descriptive statistics for each study group are presented in **Tables 2**, **3**.

In order to identify basal differences between groups (intervention/control), univariate ANOVA models were applied with the pre-selected variables as dependent (separate analysis for each variable), *group* (intervention/control) as the fixed factor; and *age*, *gender* and *socioeconomic group* (UBN/SBN) as covariables. Results showed no significant differences between intervention and control groups, for all the socioenvironmental pre-selected variables (**Table 2**).

We then evaluated the assumptions for mixed models procedures, including residual normality, homocedasticity and independence. For this purpose, descriptive and univariate analyses, histograms and plot graphics as well as Levene tests were executed for each variable. All dependent variables showed violations of at least one of the considered criteria, and therefore these variables were transformed for the analysis (using square root or arcsine transformations). Finally, for each dependent variable, scores were transformed to *z*-scores prior to their inclusion in the mixed model analyses. This was done to have a common metric to compare intervention outcome across the tasks. Means and standard deviations for each cognitive task are presented in **Table 5**. Regarding basal cognitive performance, univariate ANOVA models were executed for each dependent variable in order to compare basal performance between the study groups. Analysis included *group* (intervention/control) as the fixed factor; baseline performance variables of each task were the dependent variables (separate analysis for each cognitive process); and *age*, *gender,* and *socioeconomic group* (UBN/SBN) were the covariables. Results indicated that both study groups were homogeneous with respect to their basal cognitive performance (**Table 4**).

Considering the sample sizes and the extensive number of preselected independent variables to enter as predictors, we decided to reduce them with different procedures including: principal component analysis (PCA) from a set of variables, and correlation analysis (see next section).

Selection of potential predictors

A PCA was executed for variables selected from the Socioeconomic Status Scale and the LISRES inventory (see section Socioeconomic, Life, and Health Condition Measures) (PCA with a *Promax* rotation). The criteria used for the selection of the final PCA model were Eigenvalues over 1.00; Kaiser Coefficients over 0.6, total value of the commonalities over 10 and value of the commonalities for each variable over 0.4. The application of this procedure resulted in the identification of six factors (**Table 5**): Factor 1 (Household economic status) involves *economic* and *housing stressors*, and *economic resources*; Factor 2 (Family context) concerns *couple* and *child stressors*, *negative life events*, and *couple resources*; Factor 3 (Socioeconomic status) comprises *parental education* and *parental occupation level*, *housing conditions,* and *overcrowding*; Factor 4 (Social resources) involves *child*, *family*, and *social resources*; Factor 5 (Ties support) concerns *social* and *family stressors*; and Factor 6 (Life events)

Table 2 | Sociodemographic information of the SIP sample by group (continuous variables).

Characteristic	Intervention			Control				
	n	Mean	*SE*	*n*	Mean	*SE*	*F**	*P**
Child age (at baseline)	161	4.50	0.07	161	4.53	0.06	0.096	0.757
SOCIOECONOMIC INFORMATION[a]								
Parent education level[b]	169	6.08[c]	0.21	167	6.04[c]	0.19	0.002	0.961
Parent occupation background[b]	169	2.92[d]	0.17	170	2.89[d]	0.17	0.003	0.955
Housing[e]	166	8.64	0.14	167	8.95	0.14	2.288	0.131
Overcrowding conditions[f]	168	5.65	0.18	169	6.07	0.16	3.591	0.059
LIFE STRESSORS AND SOCIAL RESOURCES[g]								
Physical health	125	−49.38	0.88	129	−50.22	0.84	0.479	0.489
Housing stressors	127	−60.14	1.13	131	−58.79	1.12	0.728	0.394
Economic stressors	127	−65.71	0.63	131	−64.17	0.76	2.442	0.119
Working stressors	68	−48.40	0.79	69	−50.06	1.04	1.606	0.207
Couple stressors	103	−55.77	0.90	108	−56.35	0.84	0.227	0.634
Child stressors	125	−67.89	0.92	131	−66.96	0.85	0.548	0.460
Family stressors	112	−46.99	0.84	116	−47.68	1.00	0.276	0.600
Friends and social life stressors	107	−46.55	0.94	108	−45.52	0.83	0.682	0.410
Negative life events	127	−55.36	1.08	130	−54.49	1.05	0.331	0.565
Economic resources	126	38.36	0.07	130	38.50	0.09	1.484	0.224
Working resources	68	50.54	0.38	72	50.03	0.44	0.767	0.383
Couple resources	103	55.14	0.61	107	54.97	0.63	0.035	0.851
Child resources	125	65.95	0.47	130	65.94	0.52	0.000	0.985
Family resources	112	49.39	0.71	117	47.86	0.79	2.072	0.151
Friends and social life resources	105	49.96	0.98	107	47.28	1.23	2.879	0.091
Positive life events	127	48.77	0.77	130	50.74	0.77	3.274	0.072

[a] *Socioeconomic information was obtained for most cases (this is the reason for the higher sample sizes in those variables).*

[b] *Highest educational and occupational levels reached by parents.*

[c] *Incomplete secondary school level.*

[d] *Non-skilled worker.*

[e] *Scale range: 3–12 points, with higher scores for better housing conditions.*

[f] *Scale range: 0–9 points, with higher scores for better conditions.*

[g] *T-scores from each item evaluated in the Life Stressors and Social Resources Inventory (LISRES).*

**Univariate ANOVA was performed for each variable.*

Table 3 | Sociodemographic information of the SIP sample by group (categorical variables).

Characteristic	Intervention ($n = 170$)		Control ($n = 173$)	
	n	%	*n*	%
CHILD GENDER				
Male	87	51.18	90	52.02
Female	83	48.82	83	47.98
HEALTH HISTORY				
With history of medical illness[a]	5	2.94	3	1.73
Without history of medical illness	165	97.06	170	98.27
FREQUENCY OF TRAINING SESSION				
Once a week	139	81.80	120	69.40
Twice a week	31	18.20	53	30.60

[a] *Low weight at birth, premature, neurological, and/or perinatal disorders.*

involves *positive life events*. All these factors were incorporated into the data set, and for all of them, higher scores refer to better environmental conditions.

We performed a Pearson Correlation analysis including all potential predictors (socio-environmental factors derived from the PCA, and other variables not included in the PCA: demographic information, child health records and training exposure information) to identify variables with significant and high associations (Pearson coefficient over 0.5, $p < 0.05$). In those cases, only one of the correlated variables was selected for the subsequent steps—the selection was made based on the reliability of measures. The degree of association among independent and dependent variables was separately analyzed. For both, dependent and independent variables results showed no significant associations between variables (**Tables 6**, **7**).

Final models for the prediction analysis

Regarding the methodological approaches to analyze how ecological factors (i.e., micro- and mesosystemic) affect development,

Table 4 | Performance by task, time of assessment, and group in the SIP.

Time	Task	Control			Intervention				
		n	Mean	*SE*	*n*	Mean	*SE*	*F*	*p*
1 (pre-intervention)	Tower of London	144	18.59	1.58	141	16.52	1.42	1.09	0.298
	Corsi blocks	124	11.56	0.64	113	12.81	0.74	1.51	0.220
	FIST*	96	4.66	0.33	73	3.89	0.33	2.02	0.157
	Selective attention	119	0.50	0.03	112	0.52	0.03	0.09	0.765
	Day/night	148	0.75	0.02	147	0.76	0.02	1.11	0.292
2 (post-intervention)	Tower of London	121	26.71	1.71	130	43.17	1.71	56.75	0.000
	Corsi blocks	93	15.40	0.92	82	23.20	1.39	27.40	0.000
	FIST*	67	5.49	0.31	56	6.50	0.29	6.78	0.010
	Selective attention	95	0.63	0.03	83	0.77	0.02	23.45	0.000
	Day/night	122	0.83	0.02	121	0.88	0.02	4.73	0.031

*Time, moment of cognitive assessment; *this task was implemented in the second and third year of the program implementation (2003/2004).*

Table 5 | PCA results depicting variables associated with socioeconomic status and level of stressors and resources in the SIP.

	Household economic status	Family context	Socio-economic status	Social resources	Ties support	Life events
Economic stressors[a]	**0.84**	0.00	0.03	−0.12	−0.06	0.24
Economic resources[a]	**0.82**	−0.11	0.02	−0.01	−0.01	0.08
Housing stressors[a]	**0.60**	0.09	−0.03	0.12	0.16	−0.30
Couple stressors[a]	−0.02	**0.79**	0.12	0.05	0.11	0.15
Couple resources[a]	0.18	**0.64**	−0.05	0.24	−0.21	0.06
Child stressors[a]	−0.27	**0.58**	−0.01	0.11	0.29	0.14
Negative life events[a]	0.13	**0.42**	0.22	−0.20	0.24	**−0.41**
Overcrowding[b]	−0.19	0.15	**0.79**	−0.13	−0.21	−0.03
Housing conditions[b]	0.12	−0.10	**0.67**	0.05	−0.05	−0.27
Parents occupation level[b]	0.18	0.22	**0.56**	0.05	−0.05	0.09
Parents education level[b]	0.11	−0.37	**0.47**	0.26	0.24	0.22
Child resources[a]	0.06	0.10	0.03	**0.72**	−0.04	−0.02
Social resources[a]	−0.17	0.00	0.11	**0.70**	0.06	0.04
Family resources[a]	0.05	0.21	−0.15	**0.67**	−0.09	−0.15
Family stressors[a]	0.06	0.04	−0.14	−0.01	**0.75**	−0.11
Social stressors[a]	−0.04	0.08	−0.07	−0.04	**0.68**	0.09
Positive life events[a]	0.17	0.24	−0.05	−0.10	0.04	**0.87**

n = 221.

[a] Variables from the LISRES inventory.

[b] Variables from the Socioeconomic Status Scale.

Bold text indicates variables loading on each factor. These six factors were the only ones with eigenvalues larger than 1 in the correlation matrix, and the Scree plot also suggested the six-factor solution.

one of the traditional methods is the analysis of variance for repeated measures. During the past decade, a number of analytical models that overcome some disadvantages of the previous models (e.g., the ability to handle missing data) have been implemented for this type of analysis. These models are known as general linear mixed models (GLMM) (Long and Pellegrini, 2003; Singer and Willett, 2003; Arnau and Balluerka, 2004; Ferrer et al., 2004; Arnau and Bono, 2008; Seltman, 2009). Based on that, we conducted a sequence of mixed model analyses to identify significant predictors associated with higher levels of cognitive performance pre-intervention and with more improvement in cognitive performance from pre- to post-intervention.

We first conducted mixed model analyses with a basal predictor (*time*) and the interaction between *time* and *group* (intervention and control), in order to identify differences at baseline performance and trajectories (training impact) between both groups. Results showed a significant effect of *time* (*Attention*: $B = 0.916$, $p < 0.0001$; *Working memory*: $B = 1.076$, $p < 0.0001$; *Inhibitory control*: $B = 0.396$, $p = 0.0004$; *Flexibility*: $B = 0.899$, $p < 0.0001$; and *Planning*: $B = 1.219$, $p < 0.0001$), which means that all children (intervention and control) increased their

Table 6 | Correlations for independent variables in the SIP.

	1	2	3	4	5	6	7	8	9	10	11
1. Factor 1 (Household economic status)	–										
2. Factor 2 (Family context)	0.265**	–									
3. Factor 3 (Socioeconomic status)	0.272**	0.002	–								
4. Factor 4 (Social resources)	0.156*	0.106	0.093	–							
5. Factor 5 (Ties support)	0.047	0.072	0.023	0.046	–						
6. Factor 6 (Life events)	−0.05	−0.138*	0.103	0.06	−0.102	–					
7. Maternal stress for physical health problems	−0.038	0.055	−0.055	−0.038	0.133	0.057	–				
8. Child sex	−0.059	−0.109	−0.045	−0.048	0.022	−0.043	0.000	–			
9. Health records	0.075	0.046	0.165*	0.024	0.281**	−0.075	0.119*	−0.018	–		
10. Child age	−0.15	0.013	−0.142	0.161	−0.073	−0.073	0.014	0.06	−0.081	–	
11. Frequency of sessions	0.046	0.121	−0.04	0.046	0.118	0.054	0.06	0.067	0.06	0.045	–

*$p < 0.05$; **$p < 0.01$.

Table 7 | Correlations for dependent variables in the SIP.

	1	2	3	4	5
1. Planning	–				
2. Visuo-spatial organization	0.380**	–			
3. Cognitive flexibility	0.335**	0.338**	–		
4. Attentional control	0.425**	0.307**	0.375**	–	
5. Inhibitory control	0.155**	0.195**	0.208**	0.189**	–

**$p < 0.01$.

baseline performance on all tasks. Additionally, results evidenced significant effects of *group* for most of the dependent variables (*Attention*: $B = -0.493$, $p = 0.00004$; *Working memory*: $B = -0.590$, $p < 0.0001$; *Flexibility:* $B = -0.569$, $p = 0.0069$; and *Planning:* $B = -0.750$, $p < 0.0001$), which means that children in the intervention group demonstrated improved performance after training compared to children in the control group. In addition, no significant differences at baseline were identified between groups.

Second, independent variables were grouped into four blocks of information: (1) *living conditions at home, life stressors, and social resources* (including the six factors derived from the PCA analysis); (2) *demographic information* (*child age* and *gender*, and *maternal stress for physical health problems*); (3) *child health* (*health records*); and (4) *training exposure* (*frequency of sessions* and *group*). Analyses were executed separately for each block. The interactions between independent variables with *time* and *group* (intervention or control) were included in the models, in order to identify differences between both groups at baseline performance and cognitive trajectories after training.

Before the next step, we tested the missing completely at random (MARC) assumption for the independent variables included in the blocks, and the cognitive performance variables. The assumption was verified for the independent variables ($X^2 = 22.85$, $p = 0.196$), but not for the dependent cognitive variables. However, we did not input cognitive data based on the notion that doing this could alter the slope of the trajectories.

For each block, mixed model analyses were executed several times, removing the non-significant variables each time. This procedure was repeated until only significant variables were included for each block for each given cognitive outcome (dependent variable). The purpose of this was to reduce the number of independent variables to generate a final model of prediction to detect significant variables associated with cognitive performance. In general, results from this step showed significant socio-environmental predictors for each dependent variable, and overall, they evidenced a similar pattern, and yet also some differences, between the cognitive control processes and programs. A summary of results from these analyses is available from the authors upon request.

We combined the significant variables detected from each block in the previous step and included them in a final model of predictors. Similarly to what we explained above, we executed mixed model analyses several times, removing every time the non-significant predictors. At the end of this procedure, we identified a set of significant predictors (final model). This step was also executed for each dependent variable.

Each step was performed to ensure that the final model adequately reflected predictors associated with levels of pre-intervention cognitive performance and with improvement in cognitive performance from pre- to post-intervention. For the number of comparisons (*attention* = 4, *workingmemory* = 5, *inhibitorycontrol* = 2, *flexibility* = 6, *planning* = 4), the Bonferroni correction was used for a 0.05 level of significance (the final values of p were: attention = 0.0125, *workingmemory* = 0.01, *inhibitorycontrol* = 0.025, *flexibility* = 0.0083, *planning* = 0.0125).

Variables associated to cognitive performance and trajectories (final models)

We selected predictors for a final model for each program based on the results from previous steps. **Table 8** includes the final parameter estimates for each cognitive process, and show significant predictors of cognitive performance at Time 1 (baseline) and of intervention trajectories (difference between Time 2 and Time 1). It is important to point out that for the final models there was a reduction in the number of participants due to the lack of information for all the predictors for some children.

In the final model for *Attention* (Model = *time, group, age, time*group*; Pseudo $R^2 = 0.0522$; $n = 329$), the main effect

Table 8 | Results for the final model for each cognitive process in the SIP.

Dependent variable[a]	Parameters	Estimate (*SE*)	η^2
Attention ($n = 329$)	Intercept	−1.578 (0.157)***	–
	Time	0.930 (0.098)***	0.349
	Child age	0.564 (0.058)***	0.295
	Group (control)*time	−0.477 (0.135)**	0.060
Working memory ($n = 138$)	Intercept	−1.348 (0.226)***	–
	Time	0.810 (0.102)***	0.404
	Group (control)	−0.398 (0.136)*	0.078
	Child age	0.540 (0.087)***	0.278
	Social resources*time	0.202 (0.092)*	0.050
Inhibitory control ($n = 382$)	Intercept	−0.141 (0.141)***	–
	Time	0.334 (0.075)***	0.064
	Child age	0.348 (0.051)***	0.131
Flexibility ($n = 329$)	Intercept	−1.284 (0.223)***	–
	Time	1.920 (0.294)***	0.185
	Child age	0.516 (0.075)***	0.138
	Frequency (once a week)	−0.577 (0.103)***	0.163
	Group (control)*time	−0.550 (0.191)*	0.050
	Child age*time	−0.423 (0.109)**	0.085
Planning ($n = 329$)	Intercept	−2.133 (0.116)***	–
	Time	1.192 (0.078)***	0.444
	Child age	0.667 (0.039)***	0.501
	Group (control)*time	−0.720 (0.111)***	0.130

Estimates from Proc Mixed using the Restricted Maximum Likelihood estimator.
[a]Dependent Variables = Z-scores; parameter estimate standard errors (SE) listed in parentheses.
$^*p < 0.05$; $^{**}p < 0.001$; $^{***}p < 0.0001$.

of *time* shows that children from both groups, on average, significantly increased their basal performance around one standard deviation after training ($B = 0.930$; $p < 0.0001$). In addition, results show that children in the intervention group had higher performance after training than children in the control group ($B = -0.477$, $p = 0.0005$). Results also show effects of child *age* on Time 1 ($B = 0.564$; $p < 0.0001$). This result indicates that performance at baseline was higher for older children.

In the final model for *Working memory* [Model = *time, group, age, social resources (factor 4), time*social resources (factor 4)*; Pseudo $R^2 = 0.2055$; $n = 138$], the main effect of *time* shows that children from both groups increased their basal performance around one standard deviation ($B = 0.810$; $p < 0.0001$). In this final model the interaction between *time* and *group* was not included (due to dropping out of the model as non-significant; thus, differences between groups in the cognitive trajectory were not evaluated). Despite that, it is important to consider that results from previous analytical steps (prior to the inclusion of the predictors) show significant differences in cognitive trajectories between groups (i.e., more improvement in the intervention group).

With respect to the prediction of trajectories, results show that for each point on the *social resources* score (which means the perception of more resources associated with family, children and friends) children increased 0.202 points between pre and post-intervention performance in this task ($p = 0.0303$). Results also show effects of child *age* ($B = 0.540$; $p < 0.0001$) and *group* ($B = -0.398$; $p = 0.0042$) at Time 1. This pattern of results suggests that performance on Time 1 was higher for older children, and that children in the control group had, on average, lower baseline performance than children assigned to the intervention group.

For the *Inhibitory control* variable, results from the final model (Model = *time*, *age*, Pseudo $R^2 = 0.0203$; $n = 382$) show main effects of the two variables included in the model. With respect to the effect of *time*, results suggest that children in both groups, on average, increased their initial performance after the intervention ($B = 0.334$; $p < 0.0001$). As in the previous case, in this model the variable *group* was not included. Also, results from previous steps showed non-significant differences in cognitive trajectories between groups for this task (i.e., both groups had similar change in performance from pre- to post-test).

Regarding the effect of *age* ($B = 0.348$; $p < 0.0001$), results suggest that older children had higher scores on this task. Moreover, our results show that none of the socio-environmental variables were related to the change in inhibitory control from pre- to post-intervention assessment.

In the final model for *Flexibility* (Model = *time, group, age, frequency of sessions, time*group, time*age*; Pseudo $R^2 = 0.1728$; $n = 329$), the main effect of *time* shows that children in both groups increased their initial performance around two standard deviations ($B = -1.920$; $p < 0.0001$). In addition, results show that children in the intervention group had a higher increase in their performance after training than the control group ($B = -0.550$; $p = 0.0046$). Likewise, change in flexibility after training was also associated with child *age* ($B = -0.423$; $p = 0.0002$), which suggests that older children had smaller increases in their performance after training. Results also show that older children had higher scores at baseline ($B = 0.516$; $p < 0.0001$), and children who were involved in the intervention with a frequency of one session per week had 0.577 lower scores than children who were involved in the intervention with a frequency of two times per week ($p < 0.0001$).

Finally, in the model for *Planning* (Model = *time, group, age, time*group*; Pseudo $R^2 = 0.1355$; $n = 329$), the main effect of *time* shows that children from both groups increased their initial performance around one standard deviation ($B = 1.192$; $p < 0.0001$). Results also suggest significant effects of *group* on cognitive trajectories after training, such that children in the intervention group had higher scores after the intervention than children in the control group ($B = 0.720$; $p < 0.0001$). Also, our results indicate effects of child *age* ($B = 0.667$; $p < 0.0001$) on Time 1 performance, which suggest that scores tended to be higher for older children.

CTP PROGRAM

In this program we implemented the same procedures as in the SIP. First, a set of variables was pre-selected as potential predictors of cognitive performance at baseline and of the cognitive

performance change after intervention: *housing conditions, overcrowding, parental education, parental occupation, mother's physical health, housing stressors, economic stressors, working stressors, couple stressors, child stressors, family stressors, friends and social life stressors, negative life events, economic resources, working resources, couple resources, child resources, family resources, friends and social life resources, positive life events, child health records, child age, child gender, family composition, reception of social benefits (subsidies), mother age, low weight at birth, premature, neurological disorders, perinatal disorders*, and the *number of training sessions*. Descriptive statistics are presented in **Tables 9**, **10**.

We adapted some of these variables for the analysis. Specifically, *age of mother* and *number of sessions* were re-categorized into categorical predictors for the analyses in order to have comparative groups within each predictor, and be able to distinguish between very young, young or older mothers as well as low, middle or high training exposure. In both cases, based on descriptive statistics, we created three groups [1 = values below 1 standard deviation (SD) of the mean; 2 = values between 1 *SD* above and below the mean, 3 = values above 1 *SD* of the mean]. Specifically, for *age of mother* groups were: less than 24 years, between 24 and 34.4 years, more than 34.4 years. For training exposure groups were: low training exposure (less than 13 sessions), middle training exposure (13–24 sessions), and high training exposure (more than 24 sessions).

Regarding the dependent variables, we evaluated the assumptions for mixed models procedures, and, like in the SIP, all dependent variables showed violations of at least one of the criteria considered. Therefore, these variables were transformed for the analysis (using square root or arcsine transformations). Finally, for each dependent variable, scores were transformed to z to have a common metric to be able to compare intervention outcome across the tasks. Means and standard deviations for each cognitive task are presented in **Table 11**.

In this program it was also necessary to reduce the number of pre-selected variables to enter them into the analyses, and the same procedures executed in the previous program were implemented (see next section).

Table 9 | Socio-demographic information of the CTP sample (continuous variables).

Characteristic	*n*	Mean (*SD*)
Child age (at baseline)	333	3.97 (0.73)
Mother age	318	29.05 (5.76)
SOCIOECONOMIC INFORMATION[a]		
Parent education level[b]	320	7.07[c] (2.68)
Parent occupation background[b]	321	3.78[d] (2.92)
Housing[e]	321	9.91 (2.16)
Overcrowding conditions[f]	321	6.26 (2.30)
LIFE STRESSORS AND SOCIAL RESOURCES[g]		
Physical health	268	−48.97 (10.71)
Housing stressors	268	−55.49 (10.99)
Economic stressors	268	−61.87 (8.08)
Working stressors	190	−49.67 (7.64)
Couple stressors	188	−54.82 (8.88)
Child stressors	257	−64.83 (8.79)
Family stressors	258	−46.37 (8.49)
Friends and social life stressors	236	−45.82 (8.17)
Negative life events	268	−59.57 (12.96)
Economic resources	266	38.42 (0.84)
Working resources	189	51.01 (2.57)
Couple resources	189	54.34 (6.37)
Child resources	256	66.39 (5.34)
Family resources	258	48.98 (7.43)
Friends and social life resources	232	46.91 (11.12)
Positive life events	268	55.38 (10.59)
Number of training sessions[h]	292	19.34 (5.96)

[a] *Socioeconomic information was obtained in most cases (this is the reason for the higher sample sizes in those variables).*

[b] *Highest educational and occupational levels reached by parents.*

[c] *Incomplete secondary school level.*

[d] *Non-skilled worker.*

[e] *Scale range: 3–12 points, with higher scores for better housing conditions.*

[f] *Scale range: 0–9 points, with higher scores for better conditions.*

[g] *T scores from each item evaluated in the Life Stressors and Social Resources Inventory (LISRES).*

[h] *The total number of sessions vary between children due to their absence to the institutions.*

Selection of potential predictors

First, a PCA was executed for variables from the Socioeconomic Status Scale and the LISRES inventory. Same criteria used for the SIP were considered for this program, and seven components were identified (see **Table 12**): Factor 1 (Housing conditions) involves *housing conditions*, *overcrowding*, and *housing stressors*; Factor 2 (Economic status) contains *economic stressors* and *economic resources* variables; Factor 3 (Ties support) comprises *family stressors*, *negative life events* and *child stressors*; Factor 4 (Social aspects of health) concerns *social stressors*, *physical health*,

Table 10 | Socio-demographic information of the CTP sample (categorical variables).

Characteristic	*n*	%
CHILD SEX		
Male	175	52.55
Female	158	47.45
HEALTH HISTORY		
With history of medical illness[a]	160	49.54
Without history of medical illness	163	50.46
SINGLE PARENT HOUSEHOLD		
Yes	123	38.44
No	197	61.56
FAMILY WITH SOCIAL BENEFIT SUPPORT (SUBSIDIES)		
Yes	169	52.65
No	152	47.35
SOCIOECONOMIC GROUP		
Unsatisfied basic need home	161	49.80
Satisfied basic need home	160	50.20

[a] *Low weight at birth, premature, neurological, and/or perinatal disorders.*

Table 11 | Performance by task and time of assessment in the CTP.

	Time 1			Time 2		
Task	***n***	**Mean**	***SD***	***n***	**Mean**	***SD***
Tower of London	284	11.61	15.53	271	21.79	21.52
Corsi blocks	280	8.71	6.85	270	10.83	6.77
FIST	280	15.80	8.87	271	18.42	7.63
Selective attention	273	0.31	0.24	271	0.42	0.25

Time 1, cognitive assessment pre-intervention (baseline); Time 2, cognitive assessment post-intervention.

Table 12 | PCA results depicting variables associated with socioeconomic status and level of stressors and resources in the CTP.

	Housing conditions	Economic status	Ties support	Social aspects of health	Social resources	Family composition	Positive events
Housing stressors[a]	**0.83**	−0.06	0.26	0.03	−0.09	0.03	0.18
Housing conditions[b]	**0.79**	−0.05	−0.20	0.10	−0.02	0.04	−0.16
Overcrowding[b]	**0.72**	0.11	0.02	−0.29	−0.04	−0.10	0.03
Economic stressors[a]	−0.02	**0.75**	−0.03	0.18	−0.03	0.24	0.20
Economic resources[a]	0.02	**0.71**	−0.09	0.05	0.11	0.01	−0.12
Family stressors[a]	−0.01	0.01	**0.79**	0.19	−0.01	−0.12	0.05
Negative life events[a]	−0.06	0.60	**0.59**	−0.14	−0.13	−0.01	−0.23
Child stressors[a]	0.10	−0.25	**0.58**	0.26	0.10	0.09	−0.04
Social stressors[a]	−0.11	−0.03	0.18	**0.72**	−0.08	0.05	−0.04
Physical health[a]	0.00	0.24	0.16	**0.68**	−0.05	−0.02	0.14
Social benefits reception[b]	0.20	0.29	−0.08	**0.41**	0.19	−0.13	−0.35
Social resources[a]	−0.11	0.14	−0.19	−0.04	**0.76**	−0.16	0.15
Child resources[a]	−0.04	−0.22	0.19	−0.07	**0.66**	0.30	−0.16
Family resources[a]	0.02	0.12	0.54	−0.03	**0.62**	−0.04	0.09
Family composition [b]	−0.09	0.08	−0.04	0.10	−0.09	**0.88**	−0.01
Parents occupation level[b]	0.20	0.29	−0.11	−0.16	0.15	**0.63**	0.02
Positive life events[a]	0.05	−0.01	−0.03	0.07	0.09	−0.02	**0.93**

$n = 256$.

[a] *Variables from the LISRES inventory.*

[b] *Variables from the Socioeconomic Status Scale.*

Bold text indicates variables loading on each factor. These seven factors were the only ones with eigenvalues larger than 1 in the correlation matrix, and the Scree plot also suggested the seven-factor solution.

and *social benefit reception*; Factor 5 (Social resources) involves *family, child* and *social resources* variables; Factor 6 (Family composition) comprises *family composition* and *parental occupation level*; and Factor 7 (Positive events) concerns the variable *positive life events*. All these factors were incorporated into the data set, and for all of them, higher scores refer to better environmental conditions.

We performed a Pearson Correlation analysis including socio-environmental factors derived from the PCA and other variables (demographic information, child health records, and training exposure information). For the independent variables, results show a pair of variables with a moderately high degree of association: *Low weight at birth* and *Premature* ($r = 0.53$, $p < 0.0001$) (**Table 13**). Based on these results, low weight at birth was selected for the block analysis, as it can result from preterm birth or intrauterine growth restriction, or a combination of both (Shah and Ohlsson, 2002). Results show no significant correlations for the dependent variables (**Table 14**).

Creation of the final models for the prediction analysis

With the purpose of identifying significant predictors for both, basal cognitive performance and cognitive performance change between pre- to post-intervention, we conducted a sequence of mixed model analyses. We first ran a model with a basal predictor (*time*). Results showed significant estimates for all variables (*Attention*: $B = 0.4596$, $p < 0.0001$; *Working memory*: $B = 0.3816$, $p < 0.0001$; *Flexibility*: $B = 0.3458$, $p < 0.0001$; *Planning*: $B = 0.5745$, $p < 0.0001$), which means that children improved their performance on all these tasks following the group modality of cognitive training.

Second, independent variables were grouped into four blocks of information: (1) *Living conditions at home, life stressors and social resources* (including the seven factors derived from the

Table 13 | Correlations for independent variables in the CTP.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Factor 1	–															
2. Factor 2	0.359**	–														
3. Factor 3	−0.120	−0.043	–													
4. Factor 4	0.043	0.043	0.083	–												
5. Factor 5	0.270**	0.108	−0.099	0.012	–											
6. Factor 6	0.150*	0.036	0.164*	0.012	0.156*	–										
7. Factor 7	−0.092	−0.112	0.075	0.039	−0.131	−0.030	–									
8. Child sex	−0.015	−0.023	−0.094	0.035	−0.040	−0.051	−0.029	–								
9. Child age	−0.056	0.006	−0.001	0.058	−0.048	0.034	0.000	−0.067	–							
10. Parental education	0.373**	0.310**	−0.209**	−0.105	0.145*	0.129	−0.037	0.038	−0.070	–						
11. Low weight at birth	0.042	−0.081	−0.168*	−0.103	0.069	0.011	−0.044	−0.029	−0.031	0.059	–					
12. Neurological disorders	−0.006	−0.120	−0.079	−0.031	0.121	−0.008	0.035	0.030	0.073	−0.003	0.240**	–				
13. Perinatal disorders	−0.054	−0.181**	−0.026	−0.171*	0.035	0.122	−0.145*	0.028	−0.034	0.000	0.075	0.206**	–			
14. Premature	−0.023	−0.090	−0.155*	−0.076	0.015	0.104	−0.034	0.045	−0.023	0.037	0.529**	0.230**	0.189**	–		
15. Training exposure	−0.009	0.125	0.023	−0.105	−0.030	0.015	−0.051	0.006	0.047	0.030	0.025	−0.081	−0.054	0.037	–	
16. Mother age	0.076	0.050	0.082	−0.087	0.023	0.021	−0.030	−0.044	0.102	0.007	0.029	0.026	0.054	0.053	0.175**	–

*$p < 0.05$; **$p < 0.01$.

Table 14 | Correlations for dependent variables in the CTP.

	1	2	3	4
1. Planning	–			
2. Visuo-spatial organization	0.244**	–		
3. Cognitive flexibility	0.408**	0.208**	–	
4. Attentional control	0.476**	0.289**	0.486**	–

**$p < 0.01$.

PCA analysis); (2) *demographic information* (*child age* and *gender*, *parental education* and *mother age group*); (3) *child health* (*low weight at birth*, *neurological disorders*, *perinatal disorders*); and (4) *training exposure* (*training exposure group*). Analyses were executed separately for each block. The model included the interaction between each variable with *time* (see section Study Design, Participants, and Procedures).

Before the next step, the MARC assumption was tested for the independent variables included in the blocks, and the cognitive performance variables. The assumption was verified for all variables (independent variables: $X^2 = 4.46$, $p = 0.216$; basal cognitive performance variables: $X^2 = 0.931$, $p = 0.818$; post-intervention cognitive performance variables: $X^2 = 0.513$, $p = 0.916$).

For each block, we executed mixed model analyses several times, removing the non-significant variables each time. In general, results from this step showed significant socio-environmental predictors for each dependent variable, and overall, they evidenced a similar pattern, and yet also some differences between cognitive processes. As was mentioned for the SIP, the summary of results of these analyses is available upon request.

Significant variables from the previous step were combined and included in a final model of prediction (also for this step, analyses were executed several times removing every time the non-significant predictors). At the end of this procedure we identified a set of significant predictors (final model). For the number of comparisons (*attention* = 7, *workingmemory* = 3, *flexibility* = 5, *planning* = 6), the Bonferroni correction was used for a 0.05 level of significance (the final values of p were: *attention* = 0.00714, *workingmemory* = 0.01667, *flexibility* = 0.01, *planning* = 0.0083).

Variables associated to cognitive performance and trajectories (final models)

Table 15 includes the final parameter estimates for each cognitive process, and shows significant predictors of cognitive performance at Time 1 (baseline) and of intervention trajectories (difference between Time 2 and Time 1).

In the final model for Attention (Model = time, housing conditions, family composition, child age, training exposure, housing conditions*time, training exposure*time; Pseudo $R^2 = 0.0953$, $n = 188$), the main effect of time (after controlling for the other variables in the model) became non-significant ($B = 0.064$; $p = 0.7463$). However, results suggest significant effects of housing conditions ($B = 0.191$; $p = 0.003$) and marginally significant effects of training exposure (middle exposure: $B = 0.4952$; $p = 0.0217$) on cognitive trajectories. This also suggests that changes in performance from pre- to post-assessment enhanced with increasing scores on housing conditions (higher scores in this factor indicate better housing conditions, less perception of stress

Table 15 | Results for the final model for each cognitive process in the CTP.

Dependent variable[a]	Parameters	Estimate (*SE*)	η^2
Attention ($N = 188$)	Intercept	−1.621 (0.199)***	–
	Family composition	0.179 (0.051)**	0.063
	Child age	0.607 (0.070)***	0.294
	Housing conditions*time	0.191 (0.063)**	0.050
	Middle training exposure*time	0.495 (0.214)*	0.036
Working memory ($N = 215$)	Intercept	−1.429 (0.138)***	–
	Time	0.403 (0.083)***	0.105
	Ties support	0.120 (0.047)*	0.032
	Child age	0.600 (0.064)***	0.307
Flexibility ($N = 188$)	Intercept	−1.655 (0.197)***	–
	Time	0.364 (0.076)***	0.115
	Housing conditions	0.139 (0.054)*	0.037
	Family composition	0.172 (0.054)**	0.055
	Child age	0.579 (0.073)***	0.266
	Middle training exposure	0.333 (0.161)*	0.030
	High training exposure	0.430 (0.185)*	0.030
Planning ($N = 188$)	Intercept	−1.771 (0.197)***	–
	Child age	0.765 (0.071)***	0.395
	Family composition*time	0.150 (0.060)*	0.035
	High training exposure*time	0.564 (0.225)*	0.033

Estimates from Proc Mixed using the Restricted Maximum Likelihood estimator.
[a]Dependent variables = Z-scores; parameter estimate standard errors (SE) listed in parentheses.
$^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.0001$.

associated with housing, and less overcrowding conditions), and with a middle exposure to training in comparison to a low exposure. Results also show main effects of family composition ($B = 0.1791$; $p = 0.0006$) and child age ($B = 0.6067$; $p < 0.0001$) at Time 1. This pattern of results suggests that performance on the attention task was enhanced with increasing scores on family composition (higher scores for this factor indicate better parental occupation backgrounds and the presence of two parents in the home), and for older children.

In the final model for *Working memory* (Model = *time, ties support* and *child age*; Pseudo $R^2 = 0.0817$, $n = 215$), the main effect of *time* shows that children, on average, significantly increased their basal working memory performance 0.40 of a standard deviation from pre- to post-test performance ($p < 0.0001$). None of the socio-environmental variables were related to performance change in working memory between assessments. Results also show effects of *ties support* ($B = 0.1199$; $p = 0.0113$) and *child age* ($B = 0.5992$; $p < 0.0001$) on Time 1 performance. This pattern of results indicates that performance in working memory at Time 1 was higher in older children and with increasing scores on the *ties support* factor (higher scores for this factor are associated with the perception of less stress associated with the family and the children, and with less negative life events).

In the final model for *Flexibility* (Model = *time, housing conditions, family composition, child age,* and *training exposure*; Pseudo $R^2 = -0.0098$, $n = 188$), the main effect of *time* shows that children, on average, significantly increased their basal performance by 0.3641 points after training ($p < 0.0001$). None of the socio-environmental variables were related to the performance change in flexibility from pre- to post-assessment. Results also suggest significant effects of *family composition* ($B = 0.1717$; $p = 0.0016$), *child age* ($B = 0.5790$; $p < 0.0001$), and marginally significant effects of *housing conditions* ($B = 0.1386$; $p = 0.0113$) and *training exposure* (middle exposure: $B = 0.3329$, $p = 0.0402$; high exposure: $B = 0.4304$, $p = 0.0209$) at Time 1. This pattern of results indicates that performance at Time 1 was higher for older children; with increasing scores on *housing conditions* (higher scores on this factor indicate better housing conditions, less perception of stress associated with housing, and less overcrowding conditions) and *family composition* (higher scores for this factor indicate better parental occupation backgrounds and the presence of two parents at home); and with high or middle exposure to training activities.

In the final model for *Planning* (Model = *time, family composition, child age, training exposure, family composition*time, training exposure*time*; Pseudo $R^2 = 0.0105$, $n = 188$), the main effect of *time* (after controlling for the other variables in the model) became non-significant ($B = 0.2634$; $p = 0.1613$). However, results suggest significant effects of *family composition* ($B = 0.1495$; $p = 0.0130$) and marginally significant effects of *training exposure* (high exposure: $B = 0.5643$; $p = 0.0131$) on cognitive trajectories. These results suggest that change between pre- and post-training performances increases with increasing scores on *family composition* (higher scores for this factor indicate better parental occupation backgrounds and the presence of two parents at home); and with high exposure to training activities. Results also show main effects of *child age* ($B = 0.7653$; $p < 0.0001$) on Time 1, which suggests that baseline performance in the planning task was higher in older children.

BASAL PERFORMANCE COMPARISONS BETWEEN SIP AND CTP

We executed univariate ANOVA models for common variables between both programs (variables for the Socioeconomic Status Scale, the Lisres inventory and performance in attentional control, visuo spatial organization and planning), in order to compare basal cognitive performance and socio-environmental factors. The model included *program* (SIP/CTP) as the fixed factor; baseline cognitive performance and socio-demographic variables were the dependent variables (analyses were run separately for each variable); and *age*, *gender,* and *socioeconomic group* (UBN/SBN) were the covariables. Comparisons between the two programs regarding socioeconomic status and life

conditions evidence significant differences between programs in some variables: overcrowding conditions [$F_{(1-658)} = 17.83$; $p < 0.0001$], economic resources [$F_{(1-522)} = 5.14$; $p = 0.024$], couple resources [$F_{(1-399)} = 4.19$; $p = 0.041$], friends and social life resources [$F_{(1-444)} = 5.55$; $p = 0.019$], positive life events [$F_{(1-525)} = 41.85$; $p < 0.0001$], child stressors [$F_{(1-513)} = 12.58$; $p < 0.0001$], family stressors [$F_{(1-486)} = 5.50$; $p = 0.019$], and negative life events [$F_{(1-525)} = 13.67$; $p < 0.0001$]. In the other variables (housing conditions, parent education and occupation levels, mother physical health, housing stressors, economic stressors, working stressors and resources, couple stressors, friends and social life stressors, working resources, child resources, and family resources) no significant differences were found between programs. With respect to cognitive performance at Time 1, results show significant differences in attentional control [$F_{(1-500)} = 20.68$; $p < 0.0001$] and visuo-spatial organization [$F_{(1-513)} = 4.45$; $p = 0.035$]. Also, results show marginally significant differences between programs in planning basal performance [$F_{(1-565)} = 3.17$; $p = 0.076$].

DISCUSSION

The main goals of the present study were to investigate: (1) how socio-environmental factors influence baseline cognitive performance; and (2) the influence of environmental factors on cognitive trajectories (based on pre- and post-intervention assessments of attention, memory, inhibitory control, flexibility and planning). We analyzed data from two intervention programs implemented in Argentina for such objectives. Both programs have their strengths and weaknesses: the SIP included a control group, and the cognitive training module consisted of an exercising approach—same materials, different trials—, whereas the CTP did not include a control group, but the cognitive training module included pedagogic activities. Despite these advantages and limitations, results allow identifying significant predictors of both basal cognitive performance and performance changes between cognitive assessments.

Although most of the socio-environmental factors considered in the present study have often been found to be related to cognitive functioning (e.g., Brooks-Gunn and Duncan, 1997; Burchinal et al., 2000; Bradley and Corwyn, 2002; Evans, 2004; Gassman-Pines and Yoshikawa, 2006; Rhoades et al., 2011; Sarsour et al., 2011), they have rarely been simultaneously considered in training studies, so their effect on training outcome has been unclear.

Results across both intervention studies show that baseline performance of healthy preschoolers on a set of basic cognitive processes (attention, working memory, inhibitory control, flexibility, and planning), and their trajectories after training and exercising (based on pre- and post-intervention assessments) can be modulated by specific socio-environmental and individual factors. Specifically, for all cognitive processes in both programs, older children had higher baseline performance. Additionally, different variables were identified as influencing performance at baseline on attention, working memory, flexibility, and planning. Specifically, in the CTP, for attention, results show that children from dual-parent households and parents with better occupational backgrounds had higher performance at baseline. The same was verified for flexibility, where in addition, performance was higher at baseline among children with better housing conditions, as well as those who had more training sessions. Finally, in the case of working memory, our results show that baseline performance was higher for children living in homes with more ties support.

This pattern of results is in agreement with the literature on the impact of poverty on cognitive performance, suggesting that worse environmental conditions (i.e., *housing conditions, parental occupation level, family composition, social resources*) predict lower cognitive performance (e.g., Conger and Brent-Donnellan, 2007; Hackman and Farah, 2009; Lipina and Colombo, 2009; Hackman et al., 2010; Rhoades et al., 2011). In addition, results could indicate a differential sensitivity of each cognitive process to different socio-environmental factors. To investigate and examine this differential sensitivity to context, similar studies with other tasks for the same processes, as well as with samples of a wider age range (from infancy through adolescence), should be implemented. In general, the literature about poverty and cognitive development is based on a broad definition of poverty. In that sense, the identification of differential sensitivity of cognitive control processes to some environmental factors would be important to the design of interventions aimed at improving cognitive performance (Lipina et al., 2011).

With respect to cognitive trajectories from pre- to post-training, different profiles were also identified for each cognitive process and intervention program. In the SIP, for attention, working memory, flexibility, and planning, training impacts were verified (children in the intervention group had more improvement than children in the control group). Additionally, trajectories for the same tasks were predicted by some environmental factors and program characteristics. Specifically, in the case of flexibility, child *age* predicted the trajectory (older children had lower change in performance from pre- to post-test). A different pattern was verified for working memory trajectories, in which the variable *social resources* was a marginally significant predictor of change (performance change increased for children from homes with more social resources).

Results in the CTP show that *housing conditions* scores predicted the attention trajectory, indicating that change in performance from pre- to post assessment was higher for children with better home conditions, less overcrowding and fewer housing stressors. A different pattern was verified for planning trajectories, where *family composition* was the significant predictor of change. That is, change in planning performance from pre- to post-test was higher for children living in homes with two parents and with better parental occupation levels. Additionally, for both tasks (i.e., attention and planning), *training exposure* was also a marginally significant predictor of change in performance, indicating that children with more training sessions tended to have higher performances after training.

It is important to note that in the CTP, the design did not include a control group because governmental agencies did not allow researchers to do that. Nevertheless, taking into account results obtained in the SIP, designed as a randomized control program, results for the CTP show similar trends regarding the associations between better socio-environmental and individual factors (e.g., age, housing conditions, and family composition) and higher cognitive performance. Specifically, in both programs children improved their basal performance in attention, working

memory, flexibility and planning. Results also suggest, for the same dependent variables that older children present larger performance increments. Additionally, for working memory, children with more family and child resources tended to perform better at baseline (CTP) or had higher post-training improvements (SIP).

Again, results for the trajectories were not homogeneous across dependent variables and programs. These variations are consistent with other studies that indicate that not all aspects of deprivation or intervention impacts affect the relation between cognitive performance and socioeconomic status (e.g., Hoff, 2006). These differences, both between cognitive performances and between programs, could suggest differential susceptibilities of each cognitive process, as well as different patterns of cognitive integration throughout development (Garon et al., 2008).

In the present study, it is important to consider that children attending both programs had different socio-demographic characteristics, and also had different cognitive performance in the same tasks at baseline stages. Furthermore, considering the absence of significant socio-environmental predictors in both programs for some baseline performances (e.g., planning in the SIP and CTP) and for cognitive trajectories (e.g., working memory, and flexibility in the CTP; and inhibitory control in the SIP), it is necessary to consider in future analyses other socio-environmental variables that could be related to those particular cognitive processes. Additionally, it would be important to consider the administration of other cognitive measures for the same processes (Lyons and Zelazo, 2011; Bauer and Zelazo, 2013). In spite of that, it is possible to conceive that not all cognitive skills are equally susceptible to training (Jolles and Crone, 2012; Rueda et al., 2012). Therefore, although we did not verify significant predictors for some baseline performances and for some cognitive trajectories in both programs, based on the present analyses we cannot conclude that socio-environmental conditions would not predict them.

Likewise, it is important to consider differences between sample characteristics when interpreting differences in results between programs and cognitive processes [e.g., intervention programs contexts of application (prekindergarten or child cares), staff instruction, supervision and curriculum design among childcare centers in CPT; districts of implementation; and PCA results] (Ramey and Ramey, 2003; Barnett, 2011; Reynolds et al., 2011; Weiland and Yoshikawa, 2012).

Besides the need to deepen the analysis of different sensitivity to the context of cognitive processes (both, baseline performances and the effects of interventions), studies with different cognitive measures, socio-environmental variables, different levels of analysis-such as individual susceptibility or sensitive periods (Thomas and Johnson, 2008; Obradović et al., 2010), would contribute in such a sense.

Finally, it is important to mention that results must be tempered by some study limitations. First, each of the cognitive processes was measured using a single task. Future studies should consider using a variety of tasks that target the same cognitive process. Second, the analysis of change of performance over time by training is based on two time points of measurement. Although short time effects can be evaluated, future studies should include the analysis of long-term effects of training for a better understanding of the links between family and child background and training impact. Third, it is possible that the analyses in this study were underpowered, perhaps partially due to the number of evaluated subjects. Despite that, results tend to be similar to those of other studies that have used socio-environmental factors as well as training exposure to predict cognitive development (e.g., Ramey and Ramey, 2003; Rhoades et al., 2011). Finally, with respect to the program design, the CTP did not include a control group as the aim was to compare two training modalities (individual/group). Although results showed similar profiles in both programs, studies should include control designs.

Overall, this work contributes to elucidating the complex relationships between socio-environmental factors, cognitive development and intervention strategies, suggesting that environmental factors could be associated in particular ways with performance in tasks demanding attention, working memory, inhibitory control, cognitive flexibility and planning.

CONCLUSIONS

Analysis suggests that environmental factors moderated cognitive performance at baseline and through the course of the interventions in some, but not all, cognitive processes.

In sum, the contribution of the present study consists in the identification of factors that contribute to performance changes after cognitive interventions. The methodology implemented gives additional information about the impact of training, traditionally evaluated by comparing pre and post mean scores. It also contributes to the current literature about the emergence and development of cognitive processes, and their modulation by interventions in longitudinal analyses. The implemented approach and results are important for informing future intervention designs for both children and their families in Argentina.

AUTHOR CONTRIBUTIONS

M. Soledad Segretin participation as operator and supervisor in both training programs; tabulation and preparation of datasets; statistical analysis design and execution; manuscript writing. Sebastián J. Lipina participation as designer and coordinator in both intervention programs; statistical analysis design and supervision; manuscript writing. M. Julia Hermida collaboration in preparation of datasets; manuscript review. Tiffany D. Sheffield statistical analysis execution and supervision; manuscript review. Jennifer M. Nelson statistical analysis design and supervision; manuscript review. Kimberly A. Espy statistical analysis design and supervision; manuscript review. Jorge A. Colombo participation as designer in both intervention programs; manuscript review.

ACKNOWLEDGMENTS

Research was supported by UNICEF-Argentina, Fundación Bunge y Born, CONICET, Secretaría de la Niñez y de la Familia (Gobierno de Salta), Fundación Conectar, OMHSA, SAMSA, Unidad de Neurobiología Aplicada (UNA, CEMIC-CONICET); and NIH Grant 2 R01 MH065668 (Kimberly Espy). The authors thank the following people for their contribution to the programs implementations: María Inés Martelli, Beatriz Vuelta, Verónica Cristiani, Marisol Blanco, Irene Injoque-Ricle, and Beatriz Stuto.

REFERENCES

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724

Appleyard, K., Egeland, B., van Dulmen, M. H. M., and Sroufe, L. A. (2005). When more is not better: the role of cumulative risk in child behavior outcomes. *J. Child Psychol. Psychiatry* 46, 235–245. doi: 10.1111/j.1469-7610.2004.00351.x

Arnau, J., and Balluerka, N. (2004). Análisis de datos longitudinales y de curvas de crecimiento. Enfoque clásico y propuestas actuales. [longitudinal and growth trajectory data analysis. traditional approach and current proposals.] *Psicothema* 6, 156–162. Available online at: http://www.psicothema.com/psicothema.asp?ID=1175

Arnau, J., and Bono, R. (2008). Estudios longitudinales. modelos de diseño y análisis. [longitudinal studies. desing and analysis models.] *Escritos de Psicología* 2, 32–41. Available online at: http://www.redalyc.org/articulo.oa?id=271020194005

Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science* 333, 975–978. doi: 10.1126/science.1204534

Bauer, P. J., and Zelazo, P. D. (2013). Ix. Nih toolbox cognition battery (cb): summary, conclusions, and implications for cognitive development. *Monogr. Soc. Res. Child Dev.* 78, 133–46. doi: 10.1111/mono.12039

Beatty, A. (2009). *Benefit-Cost Analysis for Early Childhood Interventions*. Washington, DC: National Academies Press.

Belsky, J., Bakermans-Kranenburg, M. J., and van IJzendoorn, M. H. (2007). For better and for worse. differential susceptibility to environmental influences. *Curr. Dir. Psychol. Sci.* 16, 300–304. doi: 10.1111/j.1467-8721.2007.00525.x

Berch, D. B., Krikorian, R., and Huha, E. M. (1998). The corsi block-tapping task: methodological and theoretical considerations. *Brain Cogn.* 38, 317–338. doi: 10.1006/brcg.1998.1039

Berkman, E. T., Graham, A. M., and Fisher, P. A. (2012). Training self-control: a domain-general translational neuroscience approach. *Child Dev. Perspect.* 6, 374–384. doi: 10.1111/j.1750-8606.2012.00248.x

Bierman, K. L., Nic, R. L., Greenberg, M. T., Blair, C., and Domitrovich, C. E. (2008). Executive function and school readiness intervention: impact, moderation, and mediation in the head start REDI program. *Dev. Psychopathol.* 20, 821–843. doi: 10.1017/S0954579408000394

Bodrova, E., and Leong, D. J. (1996). *Tools of the Mind. The Vygotskian Approach to Early Childhood Education*. New Jersey, NJ: Prentice Hall.

Boltvinik, J. (1995). *Poverty Measurement Methods: An Overview*. Washington, DC: PNUD.

Bradley, R. H., and Corwyn, R. F. (2002). Socioeconomic status and child development. *Annu. Rev. Psychol.* 53, 371–399. doi: 10.1146/annurev.psych.53.100901.135233

Bradley, R. H., and Corwyn, R. F. (2005). Caring for children around the world: a view from HOME. *Int. J. Behav. Dev.* 29, 468–478. doi: 10.1177/01650250500146925

Brooks-Gunn, J., and Duncan, G. J. (1997). The effects of poverty on children. *Future Child.* 7, 55–71. doi: 10.2307/1602387

Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., and Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: sources of variation in emergent mathematic achievement. *Dev. Sci.* 14. 679–692. doi: 10.1111/j.1467-7687.2010.01012.x

Burchinal, M. R., Roberts, J., Hooper, S., and Zeisel, S. A. (2000). Cumulative risk and early cognitive development. a comparison of statistical risk models. *Dev. Psychol.* 36, 793–807. doi: 10.1037/0012-1649.36.6.793

Burger, K. (2010). How does early childhood care and education affect cognitive development? an international review of the effects of early interventions for children from different social backgrounds. *Early Child. Res. Q.* 25, 140–165. doi: 10.1016/j.ecresq.2009.11.001

Buss, C., Davis, E. P., Hobel, C. J., and Sandman, C. A. (2011). Maternal pregnancy-specific anxiety is associated with child executive function at 6–9 years age. *Stress* 14, 665–676. doi: 10.3109/10253890.2011.623250

Cadima, J., McWilliam, R. A., and Leal, T. (2010). Environmental risk factors and children's literacy skills during the transition to elementary school. *Int. J. Behav. Dev.* 34, 24–33. doi: 10.1177/0165025409345045

CEPAL-UNICEF. (2010). *Pobreza infantil en América Latina y el Caribe (Childhood poverty in Latin America and the Caribbean)*. Washington DC: Naciones Unidas.

Colombo, J. A., and Lipina, S. J. (2005) *Hacia un Programa Público de Estimulación Cognitiva Infantil. Fundamentos, Métodos y Resultados de una Experiencia de Intervención Preescolar Controlada. [Towards a Cognitive Stimulation Public Program. Foundations, Methods and Results from a Controlled Intervention in Preschoolers]*. Buenos Aires: Editorial Paidós.

Conger, R. D., and Brent-Donnellan, M. (2007). An interactionist perspective on the socioeconomic context of human development. *Annu. Rev. Psychol.* 58, 175–199. doi: 10.1146/annurev.psych.58.110405.085551

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Curr. Dir. Psychol. Sci.* 21, 335–341. doi: 10.1177/0963721412453722

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Diamond, A., Barnett, W. S., Thomas, J., and Munro, S. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Engle, P. L., Black, M. M., Behrman, J. R., Cabral de Mello, M., Gertler, P. J., Kapiriri, L., et al. (2007). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *Lancet* 369, 229–242. doi: 10.1016/S0140-6736(07)60112-3

Espinet, S. D., Anderson, J. E., and Zelazo, P. D. (2012). Reflection training improves executive function in preschool-age children: behavioral and neural effects. *Dev. Cogn. Neurosci.* 4, 3–15. doi: 10.1016/j.dcn.2012.11.009

Espy, K. A., Sheffield, T. D., Wiebe, S. A., Clark, C. A. C., and Moehr, M. J. (2011). Executive control and dimensions of problem behaviors in preschool children. *J. Child Psychol. Psychiatry* 52, 33–46. doi: 10.1111/j.1469-7610.2010.02265.x

Evans, G. W. (2004). The environment of childhood poverty. *Am. Psychol.* 59, 77–92. doi: 10.1037/0003-066X.59.2.77

Evans, G. W., Li, D., and Sepanski Whipple, S. (2013). Cumulative risk and child development. *Psychol. Bull.* 16, 688–696. doi: 10.1037/a0031808

Feldman, R., and Eidelman, A. I. (2009). Biological and environmental initial conditions shape the trajectories of cognitive and social-emotional development across the first years of life. *Dev. Sci.* 12, 194–200. doi: 10.1111/j.1467-7687.2008.00761.x

Ferrer, E., Salthouse, T. A., Stewart, W. F., and Schwartz, B. S. (2004). Modeling age and retest processes in longitudinal studies of cognitive abilities. *Psychol. Aging* 19, 243–259. doi: 10.1037/0882-7974.19.2.243

Fisher, S. E. (2006). Tangled webs: tracing the connections between genes and cognition. *Cognition* 101, 270–297. doi: 10.1016/j.cognition.2006.04.004

Flouri, E., Mavroveli, S., and Panourgia, C. (2013). The role of general cognitive ability in moderating the relation of adverse life events to emotional and behavioural problems. *Br. J. Psychol.* 104, 130–139. doi: 10.1111/j.2044-8295.2012.02106.x

Flouri, E., Tzavidis, N., and Kallis, C. (2009) Adverse life events, area socioeconomic disadvantage, and psychopathology and resilience in young children: The importance of risk factors' accumulation and protective factors' specificity. *Eur. Child Adolesc. Psychiatry* 19, 535–546. doi: 10.1007/s00787-009-0068-x

Garon, N., Bryson, S. E., and Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi: 10.1037/0033-2909.134.1.31

Gassman-Pines, A., and Yoshikawa, H. (2006). The effects of antipoverty programs on children's cumulative level of poverty-related risk. *Dev. Psychol.* 42, 981–999. doi: 10.1037/0012-1649.42.6.981

Gerstadt, C. L., Hong, J. Y., and Diamond, A. (1994). The relationship between cognition and action: performance of children 3½-7 years old on a Stroop-like day-night test. *Cognition* 53, 129–153. doi: 10.1016/0010-0277(94)90068-X

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., and Strupp, B., et al. (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet* 369, 60–70. doi: 10.1016/S0140-6736(07)60032-4

Gray, R., and McCormick, M. C. (2005). Early childhood intervention programs in the US: recent advances and future recommendations. *J. Prim. Prev.* 26, 259–275. doi: 10.1007/s10935-005-3600-x

Hackman, D. A., and Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends Cogn. Sci.* 30, 1–9. doi: 10.1016/j.tics.2008.11.003

Hackman, D., Farah, M. J., and Meany, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nat. Rev. Neurosci.* 11, 651–659. doi: 10.1038/nrn2897

Hall, J. E., Sammons, P., Sylva, K., Melhuish, E., Taggart, B., Siraj-Blatchford, I., et al. (2010). Measuring the combined risk to young children's cognitive development: An alternative to cumulative indices. *Br. J. Dev. Psychol.* 28, 219–238. doi: 10.1348/026151008X399925

Hamilton, M. (1959). The assessment of anxiety states by rating. *Br. J. Med. Psychol.* 32, 50–55. doi: 10.1111/j.2044-8341.1959.tb00467.x

Hamilton, M. (1960). A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* 23, 56–62. doi: 10.1136/jnnp.23.1.56

Hoff, E. (2006). How social contexts support and shape language development. *Dev. Rev.* 26, 55–88. doi: 10.1016/j.dr.2005.11.002

Howe, L. D., Tilling, K., Galobardes, B., Smith, G. D., Gunnell, D., and Lawlor, D. A. (2012). Socioeconomic differences in childhood growth trajectories: at what age do height inequalities emerge? *J. Epidemiol. Comm. Health* 66, 143–148. doi: 10.1136/jech.2010.113068

Hughes, C., and Graham, A. (2002). Measuring executive functions in childhood: problems and solutions? *Child Adolesc. Ment. Health* 7, 131–142. doi: 10.1111/1475-3588.00024

INDEC. (2001). *Censo Nacional. Ministerio de Economía (National Census. Ministery of Economy)*. Buenos Aires: National Institute for Statistics and Census (INDEC).

INDEC. (2006). *Incidencia de la Pobreza y de la Indigencia en Menores de 14 Años de 28 Aglomerados Urbanos (Incidence of Poverty and Indigence in Children Under 14 Years of 28 Urban Areas)*. Buenos Aires: National Institute for Statistics and Census (INDEC).

Isquith, P., Crawford, J. S., Espy, K. A., and Gioia, G. A. (2005). Assessment of executive function in preschool-aged children. *Ment. Retard. Dev. Disabil.* 11, 209–215. doi: 10.1002/mrdd.20075

Jacques, S., and Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): a measure of executive function in preschoolers. *Dev. Neuropsychol.* 20, 573–591. doi: 10.1207/S15326942DN2003_2

Jolles, D. D., and Crone, E. A. (2012). Training the developing brain: a neurocognitive perspective. *Front. Hum. Neurosci.* 6:76. doi: 10.3389/fnhum.2012.00076

Karoly, L. A., Kilburn, M. R., and Cannon, J. S. (2005). *Early childhood interventions. Proven Results, Future Promise*. Pennsylvania, PA: RAND Labor and Population.

Kiernan, K. E., and Huerta, M. C. (2008). Economic deprivation, maternal depression, parenting and children's cognitive and emotional development in early childhood. *Br. J. Sociol.* 59, 783–806. doi: 10.1111/j.1468-4446.2008.00219.x

Kiernan, K. E., and Mensah, F. K. (2009). Poverty, maternal depression, family status and children's cognitive and behavioural development in early childhood: a longitudinal study. *J. Soc. Policy* 38, 569–588. doi: 10.1017/S0047279409003250

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., et al. (2005). Computerized training of working memory in children with ADHD- a randomized controlled trial. *J. Acad. Child Adolesc. Psychiatry* 44, 177–186. doi: 10.1097/00004583-200502000-00010

Korkman, M., Kirk, U., and Kemp, S. (1998). *NEPSY: A Developmental Neuropsychological Assessment, San* Antonio, TX: Psychological Corporation.

Leinonen, J. A., Solantaus, T. S., and Punamäki, R. L. (2002). The specific mediating paths between economic hardship and the quality of parenting. *Int. J. Behav. Dev.* 26, 423–435. doi: 10.1080/01650250143000364

Lipina, S. J., and Colombo, J. A. (2009). *Poverty and Brain Development During Childhood: An Approach From Cognitive Psychology and Neuroscience*. Washington, DC: American Psychological Association. doi: 10.1037/11879-000

Lipina, S. J., Martelli, M. I., Vuelta, B. L., and Colombo, J. A. (2005). Performance on the AnoB task of argentinean infants from unsatisfied basic needs homes. *Int. J. Psychol.* 39, 49–60. Available online at: http://www.psicorip.org/Resumos/PerP/RIP/RIP036a0/RIP03906.pdf

Lipina, S. J., Martelli, M. I., Vuelta, B. L., Injoque Ricle, I., and Colombo, J. A. (2004). Pobreza y desempeño ejecutivo en alumnos preescolares de la ciudad de buenos aires (Argentina). *Interdisciplinaria* 21, 153–193. Available online at: http://www.scielo.org.ar/pdf/interd/v21n2/v21n2a02

Lipina, S. J., and Segretin, M. S. (2006). "Pobreza y salud mental infantil: aportes de la neurociencia cognitiva," en *Psicología de la Salud Infantil, Capítulo 11,* eds M. Osorio and L. A. Oblitas. ["Poverty and mental health in children: contribution from the cognitive neuroscience," in *Child Health Psychology, Chapter 11*] (Bogotá: Psicom Editores), 1–24.

Lipina, S. J., Segretin, M. S., Hermida, M. J., and Colombo, J. A. (2012). "Research on child poverty and development from a cognitive neuroscience perspective: examples of studies in argentina," in *Child and Adolescent Mental Health, Chapter 17* ed U. S. Nayar (Mumbai: SAGE), 256–274.

Lipina, S. J., Simonds, J., and Segretin, M. S. (2011). Recognizing the child in child poverty. *Vulnerable Child. Youth Stud.* 6, 8–17. doi: 10.1080/17450128.2010.521598

Lipina, S., Segretin, S., Hermida, J., Prats, L., Fracchia, C., López Camelo, J., et al. (2013). Linking childhood poverty and cognition: environmental mediators of non-verbal executive control in an argentine sample. *Dev. Sci.* 16, 697–707. doi: 10.1111/desc.12080

Long, J. D., and Pellegrini, A. D. (2003). Studying change in dominance and bulling with linear mixed models. *School Psych. Rev.* 32, 401–417.

Lyons, K. E., and Zelazo, P.D., (2011). Monitoring, metacognition, and executive function: elucidating the role of self-reflection in the development of self-regulation. *Adv. Child Dev. Behav.* 40, 379–412. doi: 10.1016/B978-0-12-386491-8.00010-4

Marsh, H. L., Stavropoulos, J., Nienhuis, T., and Legerstee, M. (2010). Six- and 9-month-old infants discriminate between goals despite similar action patterns. *Infancy* 15, 94–106. doi: 10.1111/j.1532-7078.2009.00002.x

Martelli, M. I., Vuelta, B. L., Blanco, M., Cristiani, V., Segretin, M. S., Lipina, S. J., et al. (2007). "Programas de intervención temprana en nuestro país. Experiencia de una aplicación individual de estimulación cognitiva," en *Pobreza y desarrollo infantil. Una contribución multidisciplinaria, Capítulo 10,* ed J. A. Colombo, ["Early intervention programs in our country. Experience of an individual application of cognitive stimulation," in *Poverty and child development. A multidisciplinary contribution, Chapter 10*] (Buenos Aires: Editorial Paidós), 229–250.

McCandliss, B. D., Beck, I. L., Sandak, R., and Perfetti, C. (2003). Focusing attention on decoding for children with poor reading skills: design and preliminary tests of the word building intervention. *Sci. Stud. Read.* 7, 75–104. doi: 10.1207/S1532799XSSR0701_05

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *Am. Psychol.* 53, 185–204. doi: 10.1037/0003-066X.53.2.185

Meunier, J. C., Boyle, M., O'Connor, T. G., and Jenkins, J. M. (2013). Multilevel mediation: Cumulative contextual risk, maternal differential treatment, and children's behavior within families. *Child Dev.* 84, 1594–1615. doi: 10.1111/cdev.12066

Mikulic, I. M. (1999). *La evaluación Psicológica de los Recursos Sociales y los Estresores de vida. Aportes del Inventario LISRES. [Psychological Assessment of Social Resources and Life Stressors. Contributions from the LISRES Inventory].* Buenos Aires: Saint Claire Editora.

Najman, J. M., Aird, R., Bor, W., O'Callaghan, M., Williams, G. M., and Shuttlewood, G. J. (2004). The generational transmission of socioeconomic inequalities in child cognitive development and emotional health. *Soc. Sci. Med.* 58, 1147–1158. doi: 10.1016/S0277-9536(03)00286-7

Obradović, J., Bush, N. R., Stamperdahl J., Adler, N. E., and Boyce, W. T. (2010). Biological sensitivity to context: the interactive effects of stress reactivity and family adversity on socioemotional behavior and school readiness. *Child Dev.* 81, 270–289. doi: 10.1111/j.1467-8624.2009.01394.x

Perez-Johnson, I., and Maynard, R. (2007). The case for early, targeted interventions to prevent academic failure. *Peabody J. Educ.* 82, 587–616. doi: 10.1080/01619560701602983

Ramey, C. T., and Ramey, S. L. (1998). Early intervention and early experience. *Am. Psychol.* 53, 109–120. doi: 10.1037/0003-066X.53.2.109

Ramey, S., and Ramey, C. T. (2003). "Understanding efficacy of early educational programs: critical design, practice, and policy issues," in *Early Childhood Programs for a New Century,* eds A. J. Reynolds, M. C. Wang, and H. J. Walberg (Washington, DC: Child Welfare League of America Press), 35–65.

Raver, C. C., Blair, C., and Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: new perspectives on models of differential susceptibility. *Dev. Psychol.* 49, 292–304. doi: 10.1037/a0028343

Raver, C. C., Gershoff, E. T., and Aber, J. L. (2007). Testing equivalence of mediating models of income, parenting, and school readiness for white, black, and hispanic children in a national sample. *Child Dev.* 78, 96–115. doi: 10.1111/j.1467-8624.2007.00987.x

Reynolds, A. J., and Temple, J. A. (2005). Priorities for a new century of early childhood programs. *Infants Young Child.* 18, 104–118. doi: 10.1097/00001163-200504000-00004

Rhoades, B. L., Greenberg, M. T., Lanza, S. T., and Blair, C. (2011). Demographic and familial predictors of early executive function development: contribution of a person-centered perspective. *J. Exp. Child Psychol.* 38, 638–662. doi: 10.1016/j.jecp.2010.08.004

Rueda, M. R., Cómbita, L. M., Paz-Alonso, P. M., Pozuelos, J. P., and Abundis, A. (2012). Genetic variation modulates the effect of executive attention

training in preschool children. *Cogn. Neurosci. Soc. Abstr.* D25, 109. doi: 10.1016/j.dcn.2011.09.004

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., and Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proc. Nati. Acad. Sci. U.S.A.* 102, 14931–14936. doi: 10.1073/pnas.0506897102

Sameroff, A. J., Seifer, R., Baldwin, A., and Baldwin, C. (1993). Stability of intelligence from preschool to adolescence: the influence of social and family risk factor. *Child Dev.* 64, 80–97. doi: 10.2307/1131438

Santos, D. N., O. Assis, A. M., Bastos, A. C. S., Santos, L. M., Santos, C. A. S. T., Strina, A., et al. (2008). Determinants of cognitive function in childhood: A cohort study in a middle income context. *BMC Public Health* 8:202. doi: 10.1186/1471-2458-8-202

Sarsour, K., Sheridan, M., Jutte, D., Nuru-Jeter, A., Hinshaw, S., and Boyce, W. T. (2011). Family socioeconomic status and child executive functions: the roles of language, home environment, and single parenthood. *J. Int. Neuropsychol. Soc.* 17, 120–132. doi: 10.1017/S1355617710001335

Segretin, M. S., Lipina, S. J., and Colombo, J. A. (2007a). "Intervenciones para niños en riesgo social: análisis de predicción de mejora cognitiva," en *Avances en Investigación en ciencias del comportamiento en Argentina, Capítulo 13*, ed Asociación Argentina de Ciencias del Comportamiento, ["Interventions for children at social risk: predicting cognitive imporvement," in *Progress in behavioral science research in Argentina, Chapter 13*] (Mendoza: Editorial de la Universidad Aconcagua).

Segretin, M. S., Lipina, S. J., and Petetta, D. R. (2009). Methodological and conceptual considerations for the analysis of school performance predictions, based on home and school context indexes. *Revista Iberoamericana de Evaluación Educativa* 2, 104–123. Available online at: http://rinace.net/riee/numeros/vol2-num2/art6.pdf

Segretin, M. S., Martelli, M. I., Cristiani, V. A., Blanco, M., Vuelta, B., Lipina, S. J., et al. (2007b). "Programas de intervención temprana en nuestro país. Experiencia de aplicaciones grupales de estimulación cognitiva," en *Pobreza y Desarrollo Infantil. Una Contribución Multidisciplinaria Capítulo 11,* ed J. A. Colombo, ["Early intervention programs in our country. Experience of group application of cognitive stimulation," in *Poverty and Child Development. A Multidisciplinary Contribution, Chapter 11*] (Buenos Aires: Editorial Paidós), 251–269.

Seltman, H. J. (2009). "Mixed models. A flexible approach to correlated data," in *Experimental Design and Analysis*, *Chapter 15*. Available online at: http://www.stat.cmu.edu/_hseltman/309/Book/Book.pdf

Shah, P., and Ohlsson, A. (2002). *Literature Review of Low Birth Weight, Including Small for Gestational Age and Preterm Birth.* Ontario, CA: Toronto Public Health, Toronto.

Shallice, T. (1982). Specific impairments of planning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 298, 199–209. doi: 10.1098/rstb.1982.0082

Singer, J. D., and Willett, J. B. (2003). *Applied Longitudinal Data analysis. Modeling Change and Event Occurrence.* New York, NY: Oxford Uneversity Press. doi: 10.1093/acprof:oso/9780195152968.001.0001

Stanton-Chapman, T. L., Chapman, D. A., Kaiser, A. P., and Hancock, T. B. (2004). Cumulative risk and low-income children's language development. *Topics Early Child. Spec. Educ.* 24, 227–237. doi: 10.1177/02711214040240040401

Sternberg, R. J., and Grigorenko, E. L. (2002). *Dynamic Testing.* Cambridge: Cambridge University Press.

Stevens, C., Fanning, J., Coch, D., Sanders, L., and Neville, H. J. (2008). Neural mechanisms of selective auditory attention are enhanced by computerized training: Electrophysiological evidence from language-impaired and typically developing children. *Brain Res.* 1205, 55–69. doi: 10.1016/j.brainres.2007.10.108

Temple, E., Gayle, K., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *PNAS* 100, 2860–2865. doi: 10.1073/pnas.0030098100

Reynolds, A. J., Temple, J. A., White B.A., Ou, S. R., and Robertson, D. L. (2011). Age 26 cost-benefit analysis of the child-parent center early education program. *Child Dev.* 82, 379–404. doi: 10.1111/j.1467-8624.2010.01563.x

Thomas, M. S., and Johnson, M. H. (2008). New advances in understanding sensitive periods in brain development. *Curr. Dir. Psychol. Sci.* 17, 1–5. doi: 10.1111/j.1467-8721.2008.00537.x

Thorell, L. B., Lindqvist, S., Nutley, S. B., Bohlin, G., and Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12, 106–113. doi: 10.1111/j.1467-7687.2008.00745.x

Tuñón, I. (2011). *Infancias en Observación: Progresos, Ausencias y Desigualdades. Evolución de Indicadores de Derechos y Desarrollo Humano (Childhood Observation: Progress, Absences and Inequalities. Rights and Human Development Indicator's Evolution).* Buenos Aires: Pontificia Universidad Católica Argentina, Barómetro de la deuda social de la infancia.

Walker, S. P., Wachs, T. D., Meeks Gardner, J., Lozoff, B., Wasserman, G. A., Pollitt, E., et al. The International Child Development Steering Group. (2007). Child development: risk factors for adverse outcomes in developing countries. *Lancet*, 369, 145–157. doi: 10.1016/S0140-6736(07)60076-2

Weiland, C., and Yoshikawa, H. (2012). The effects of large-scale economic change and policies on children's developmental contexts and developmental outcomes. *Child Dev. Perspect.* 6, 342–350. doi: 10.1111/j.1750-8606.2011.00222.x

Willoughby, M. T., Wirth, R. J., Blair, C. B., and The Family Life Project Investigators. (2012). Executive function in early childhood: longitudinal measurement invariance and developmental change. *Psychol. Assess.* 24, 418–431 doi: 10.1037/a0025779

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Coehn, L., and Cohen, D. (2006). Principles underlying the design of the number race, an adaptive computer game for remediation of dyscalculia. *Behav. Brain Funct.* 2, 1–14. doi: 10.1186/1744-9081-2-1

Zelazo, P. D., Carter, A., Reznick, J. S., and Frye, D. (1997). Early development of executive function: a problem-solving framework. *Rev. Gen. Psychol.* 1, 198–226. doi: 10.1037/1089-2680.1.2.198

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 September 2013; accepted: 23 February 2014; published online: 13 March 2014.

*Citation: Segretin MS, Lipina SJ, Hermida MJ, Sheffield TD, Nelson JM, Espy KA and Colombo JA (2014) Predictors of cognitive enhancement after training in preschoolers from diverse socioeconomic backgrounds. Front. Psychol. **5**:205. doi: 10.3389/fpsyg.2014.00205*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Segretin, Lipina, Hermida, Sheffield, Nelson, Espy and Colombo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 20 May 2014
doi: 10.3389/fpsyg.2014.00457

The potential adverse effect of energy drinks on executive functions in early adolescence

Tamara Van Batenburg-Eddes[1], Nikki C. Lee[1], Wouter D. Weeda[1,2], Lydia Krabbendam[1] and Mariette Huizinga[1]*

[1] *Department of Educational Neuroscience and LEARN! Research Institute, Faculty of Psychology and Education, VU University Amsterdam, Amsterdam, Netherlands*
[2] *Department of Clinical Neuropsychology, Faculty of Psychology and Education, VU University Amsterdam, Amsterdam, Netherlands*

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Kristen Mackiewicz Seghete, Oregon Health & Science University, USA
Pascal Wilhelm, University of Twente, Netherlands

****Correspondence:***
Mariette Huizinga, Department of Educational Neuroscience and LEARN! Research Institute, Faculty of Psychology and Education, VU University Amsterdam, Van der Boechorstraat 1, 1081 BT, Amsterdam, Netherlands
e-mail: m.huizinga@vu.nl

Introduction: Manufacturers of energy drinks (EDs) claim their products improve cognitive performance. Young adolescents are in a critical developmental phase. The impact of ED intake on their development is not yet clear. Therefore, we studied the associations of both caffeine intake and ED consumption with executive functions (EFs), and the role of pubertal status and sleeping problems.

Methods: A sample of 509 participants (mean age: 13.1 years, *SD* 0.85; age range: 11–16 years) participated in the study. The level of pubertal development was classified in five pubertal status categories. Participants were asked to report their caffeine (for example coffee) and ED consumption for each day of the week. In addition, they indicated sleep quality by reporting problems falling asleep or waking up and/or interrupted sleep. EFs were assessed by self- and parent reports of the Behavior Rating Inventory of Executive Function (BRIEF).

Results: Consuming on average one or more ED(s) a day was associated with more problems in self-reported behavior regulation and metacognition, and with more problems in parent-reported metacognition. Only high caffeine consumption (two or more cups a day) was associated with parent-reported problems with metacognition. The sum of caffeine and ED use was associated with a higher amount of problems with self-reported metacognition and parent reported behavior regulation. The effect estimates for the association between caffeine and ED use combined and EFs did not exceed those of EDs or caffeine separately. Adjusting for pubertal status, gender, educational level, number of sleeping problems and hours of sleep did not change the effect estimates substantially.

Conclusion: The observed associations between ED consumption and EFs suggest that regular consumption of EDs—even in moderate amounts—may have a negative impact on daily life behaviors related to EF in young adolescents.

Keywords: energy drink use, puberty, executive functions, cognitive functioning, pubertal brain development

INTRODUCTION

Since the introduction of "Red Bull" in Austria in 1987, the consumption of caffeinated drinks has grown immensely (Reissig et al., 2009). Most of these so-called "energy drinks" (EDs) are marketed directly to children and adolescents and, at the same time, the use of these drinks within this population has risen exponentially (Bramstedt, 2007). According to Seifert et al. (2011), 30 to 50% of adolescents and young adults reported in surveys to consume EDs. A recent European study showed that the prevalence of "high chronic" ED users (i.e., respondents who regularly consumed ED "4–5 days a week" or more), is highest among Dutch adolescents (27%) compared to the prevalence of "high chronic ED use" in the other participating European countries (range 7–19%; Zucconi et al., 2013). These observations suggest that prolonged and habitual use of EDs is present in a substantial group of young adolescents.

Caffeine is the most important ingredient of EDs and is typically used for its arousing effect on the central nervous system (Seifert et al., 2011). Although caffeine is a psychoactive substance, it is considered safe by the Food and Drug Administration (Temple, 2009). However, excessive use of caffeine can have detrimental health effects (Seifert et al., 2011).

In the US, EDs are classified as dietary supplements and therefore the amount of caffeine content in these drinks is not regulated [US Food and Drug Administration. Q & A on Dietary Supplements (http://www.fda.gov/Food/DietarySupplements/QADietarySupplements/default.htm)]. In the Netherlands, however, by law the maximum allowed caffeine content is 350 mg/l, which is determined by the Netherlands Food and Consumer Product Safety Authority (http://www.vwa.nl/actueel/bestanden/bestand/42527). If the caffeine content exceeds 150 mg/l, manufactures are obliged to print "high caffeine content" on their products. Furthermore, the Dutch Food Foundation advises young adolescents between the age of 13 and 18 years to consume at most one can of ED (250 ml) a day (Netherlands Food and Consumer Product

Safety Authority; Stichting Voedingscentrum Nederland, 2013).

Manufacturers of EDs claim their products improve physical and cognitive performance. The direct short-term positive effect on cognitive performance is still controversial, but when it is found it is often attributed to caffeine. The effect of caffeine on particular aspects of cognitive functioning has been observed in numerous well-controlled studies in a range of populations (Lieberman, 2001). In addition, Wesnes et al. (2013) conclude that many studies have detected improvements of cognitive functioning or alertness after ingesting caffeine or EDs. However, they also point out that most studies have investigated the short-term effects 1–2 h after ingestion. Wesnes et al. (2013) studied the effect of a specific ED 6 h after ingestion. They found a sustained effect of this specific ED in partially sleep deprived participants on four out of six cognitive performance tasks after consuming this ED compared to the placebo group (Wesnes et al., 2013). In contrast, Curry and Stasio (2009) investigated the effects of EDs alone and combined with alcohol on neuropsychological functioning. In the ED only group, they found a trend toward improved attention and no overall improvement in neuropsychological functioning from pre-to post-test. Furthermore, as little is known about the effects of EDs on cognitive functioning in young persons, Wilhelm et al. (2013) studied young adolescents aged 15 to 18 years. In a quasi-experimental design comparing three groups, no significant differences were observed between the groups that could be ascribed to the effect of ED on measures of cognitive functioning such as attention, learning ability and vocabulary. Nevertheless, all these studies focused on relatively short-term effects of EDs (i.e., effects on cognitive functioning within a certain time after consuming the ED). The effect of prolonged and habitual use of EDs on more long-term every day cognitive functioning has, to the best of our knowledge, not been studied in young adolescents.

The long-term effects of caffeine consumption and ED use during adolescence may have consequences for adolescent development. Adolescence is a period characterized by continued structural and functional brain development, triggered by the hormonal changes at the onset of puberty (Giedd, 2004; Gogtay et al., 2004; Paus, 2013). The prefrontal cortex, one of the areas of the brain that shows the greatest development during this period, contains areas involved in a variety of cognitive abilities, including executive functions. These are vital for an individual to be able to control and reflect on their behavior, and to be able to behave in a goal-directed manner. Executive functions continue to develop and improve steadily throughout adolescence and into adulthood (Huizinga et al., 2006). As young adolescents are in a critical developmental phase, this may make them in particularly vulnerable to the potential negative effects of caffeinated drinks.

Regular caffeine consumption has been related to numerous potential adverse outcomes, such as cardiovascular effects, caloric intake, diabetes and problems related to sleep (Roehrs and Roth, 2008; Seifert et al., 2011). The relation between regular caffeine consumption and disrupted sleep and increased daytime sleepiness seems to be well-established. Although adolescents seem to consume caffeinated drinks (including coffee) to a lesser extent than adults, in this age group caffeine use is also associated with sleeping problems and daytime sleepiness (see for a review: Roehrs and Roth, 2008). Furthermore, sleep seems to be particularly important during periods of brain maturation, such as adolescence (Dahl and Lewin, 2002). In addition, sleep deprivation during adolescence is related to a wide range of behavioral deficits, such as attention problems, oppositionality/irritability, behavior regulation problems, and reduced metacognitive skills (Beebe et al., 2008; O'Brien, 2009; Jackson et al., 2013).

The main goal of the present study was to investigate, in a sample of young adolescents, the associations between caffeine intake and ED consumption, and behavioral executive function and metacognition. We examined the effects of these caffeine and EDs individually and cumulatively. Because of the previously reported relations between caffeine use and sleep, and between sleep and cognitive functioning, indicators of sleeping problems were included in the current study to investigate their potential mediating effect in the relation between caffeine intake and ED consumption with cognitive functioning.

MATERIALS AND METHODS

STUDY POPULATION

This study included 564 young adolescents (M age 13.10 years, $SD = 0.85$; age range: 11–16 years; 244 females). Participants were recruited through four regular schools in urban and suburban areas in the Netherlands (see **Table 1** for sample characteristics). This study is part of a longitudinal project focusing on young adolescents' socio-emotional and cognitive development. Participants completed multiple questionnaires and cognitive tasks, including a questionnaire on executive functioning (self-report and parent report); 509 participants (and 317 of their parents) filled in this questionnaire and were therefore included in the analysis. The total sample thus consisted of 509 participants. Informed consent was obtained, and the study was approved by the Ethical Committee of the Faculty of Behavioral and Social Sciences of the University of Amsterdam. Participants did not receive credit individually, but received a voucher for an excursion together with participating classmates.

When comparing the included sample of parents ($n = 317$) to the sample of parents that was excluded because of missing parent reported data on executive functioning ($n = 192$), we found no statistically significant differences in gender, pubertal status category, caffeine or energy drink intake, or on the variables regarding sleeping, age at assessment, and BRIEF scores.

MEASURES

Caffeine and energy drink intake

Caffeine and energy drink intake were measured separately by asking the participants how often, during a normal/average week they consumed caffeine (coffee or cola), and how often they consumed energy drinks (Red Bull, Xii etc.; Graham et al., 1984; Ames et al., 2007). For each day of the week, participants indicated the number of consumed cups, cans, or glasses. These numbers were summed for caffeine use and ED intake, and for each were divided by 7 to derive the average number of consumptions of caffeine and the average number of EDs per day. As caffeine use and ED intake were correlated (Pearson $r = 0.36$, $p < 0.001$), we also calculated the total number of

Table 1 | Sample characteristics.

	Sample of $n = 509$ pupils
Gender, boys, %	52
Age at assessment in years, mean (*SD*)	13.10 (0.85)
Pubertal status category, %	
Prepubertal	8
Early puberty	23
Midpubertal	37
Late puberty	29
Postpuberty	2
Educational track, %	
Primary school	22
Pre-vocational secondary education	32
Pre-vocational/senior general secondary education	8
Senior general secondary education	10
Senior general secondary/pre-university education	16
Pre-university education	11
Caffeine, %	
<1 per day	72
≥1–2 each day	17
≥2 each day	11
EDs, %	
<1 per day	94
≥1 each day	6
Caffeine and EDs, %	
<1 per day	65
≥1–2 each day	21
≥2 each day	14
Problems falling asleep, yes, %	23
Problems staying asleep, yes, %	6
Problems waking up, yes, %	23
Total of sleeping problems, mean (*SD*)	0.51 (0.79)
Hours of sleep, mean (*SD*)	8.83 (1.26)

caffeine and ED consumptions, by summing the number of caffeine and ED consumptions. Scores were divided by 7, yielding the average number of caffeine containing drinks consumed per day.

Executive functions

Executive functions were assessed with the self-report and parent-report of the Dutch Behavior Rating Inventory of Executive Function (BRIEF: Gioia et al., 2000, 2002; Smidts and Huizinga, 2009). In contrast to experiments that enable researchers to measure specific cognitive functions as the settings can be controlled to a large degree, self-reports and parents measure executive functioning in real life settings, and thus offer increased ecological validity. The self- and parent reports of the BRIEF consists of 70 items, whereas the parent-report of the BRIEF consists of 75 items. Each item pertains to specific everyday behavior, relevant to executive functioning. Children and their parents were asked to indicate how often they or their child displayed a given problem behavior in the past 6 months. Scoring options were "1 = Never," "2 = Sometimes," or "3 = Often." Higher scores indicate more problems. Self-report items are categorized into eight clinical scales: Inhibit, Shift, Emotional Control, Monitor, Working Memory, Plan/Organize, Organization of Materials, and Persistence. In the current sample, the range of alpha's for internal consistency for the clinical subscales was 0.71 and 0.84. Parent-report items are also categorized into eight clinical scales but without the Persistence scale and with an additional scale measuring Initiative. In the current sample, the range of alpha's for internal consistency for the clinical subscales was 0.80 and 0.89. To calculate the clinical scale scores, the appropriate item scores were summed and divided by the number of items in each scale.

Two indices—the Behavior Regulation Index (BRI) and the Metacognition Index (MI)—can be formed by combining scales. The BRI represents the ability to shift cognitive set and modulate behavior and emotions, whereas the MI represents the ability to plan, organize, initiate, and hold information in mind for future-oriented problem solving. The mean across the appropriate clinical scale scores was calculated to yield the BRI and MI indices. In the current sample, the range of alpha's for internal consistency for the self-reported and parent reported indices was 0.91 and 0.96. In addition, raw scores were transformed into T-scores based on the Dutch norm population (Huizinga and Smidts, 2012).

Covariates

Pubertal status category was determined by means of the self-report Pubertal Development Scale, which was developed by Carskadon and Acebo (1993) as an adaptation of the interview-based puberty-rating scale by Petersen et al. (1988). The scale measures pubertal status using a 5-point scale to rate five questions indexing physical development. Both boys and girls are asked to rate their development with regards to growth in height, body hair growth, skin changes. For boys there are additional questions about voice change and facial hair growth and for girls there are additional questions about breast development and menarche. Answers are rated on a 4-point scale, with 1 indicating no development, and 4 indicating that development is finished, and 5 indicating "I don't know" or a missing value. The question about menarche was coded dichotomously (1 = premenarcheal, 4 = postmenarcheal). An individual's level of development was classified in terms of five pubertal status categories: prepubertal, early pubertal, mid-pubertal, late pubertal, and postpubertal. For boys, the assignment was made on the basis of reported level of body hair growth, facial hair growth, and voice change. Girls were assigned on the basis of reported level of body hair growth and breast development and whether or not a girl reported having experienced menarche (Carskadon and Acebo, 1993).

Participants answered three questions about sleep problems: (1) problems falling asleep, (2) problems staying asleep, and (3) problems waking up in the morning. Answering categories were "yes" (1), or "no" (0). The answers across the three questions about sleeping were summed yielding "total number of sleeping problems".

The demographic variables gender and educational track were also measured using a questionnaire.

STATISTICAL ANALYSES

Linear regression analysis was used to assess the associations between caffeine and/or ED intake and EF. In these analyses, caffeine and/or ED intake variables were the independent variables, all BRIEF measures were used as dependent variables. For each pair of independent and dependent variable, a separate linear regression analysis was done. Indices of behavioral regulation and metacognition were regarded as the main overall outcome measures, as these are summary measures of the clinical subscales. The associations between each independent variable and each clinical subscale measure were inspected to determine which association(s) contributed to the overall association between determinant and index. Caffeine and/or ED intake variables were dummy-coded with on average using less than one consumption of these drinks a day was used as the reference category. First, the associations between caffeine and/or ED intake and EF were adjusted for gender and pubertal status category, as gender was associated with parent reported MI [$F_{(1,315)} = 8.63$, $p = 0.004$] and pubertal status category was associated with self-reported MI [$F_{(4,438)} = 3.08$, $p = 0.016$]. Also, educational level was added as a potential confounder, as it was weakly associated with self-reported metacognition ($r = 0.16, p < 0.001$). Next, we added total number of sleeping problems and hours of sleep as potential mediating factors in the associations between caffeine and/or ED intake and EFs. The effect of sleeping problems and hours of sleep on the associations under study was small. Sleeping problems and hours of sleep were still included in our analyses because of their well-known associations with caffeine use and/or ED intake and EFs, but regression models excluding sleeping problems and hours of sleep are not presented separately. Thus, the final models were adjusted for gender, pubertal status category, educational track, sleeping problems and hours of sleep. We adjusted for multiple testing as follows. First, the *p*-value of the associations between caffeine use and/or ED intake and the overall indices was set 0.025. For this set of associations, we ran four tests. As the indices are related to each other (Pearson $r_{SR\ BRI\ \&\ MI} = 0.75$, $p < 0.001$; Pearson $r_{PR\ BRI\ \&\ MI} = 0.65$, $p < 0.001$) and the independent variables, caffeine use and ED intake, are also related (Pearson $r = 0.36$, $p < 0.001$), the desirable *p*-value 0.05 was divided by four and then multiplied by two to compensate for the correlational structure (Bender and Lange, 2001). Next, we set the *p*-value for the analyses using the clinical subscales as outcome measures at = 0.01 (i.e., 0.05 divided by five which is the maximum number of clinical subscales for the indices).

The main analyses assessing the associations between caffeine and ED intake and EFs were repeated using *T*-scores of the BRIEF scales and indices. Results were consistent with those based on the raw scores and are therefore not reported.

RESULTS

Sample characteristics are displayed in **Table 1**. Most of the participants were in the early pubertal stage (23%), midpubertal stage (37%), or late pubertal stage (29%). Eight percent of participants was in the prepubertal stage, whereas only 2% was in the postpubertal stage. 11% reported to drink, during normal weeks, on average at least two caffeine containing drinks a day such as coffee or cola (excluding energy drinks). Six percent reported to consume on average at least one energy drink a day. Problems with falling asleep and waking up were reported most often (23%). These characteristics were highly similar in the sample for which we also had parent reports of EFs.

The adjusted associations between caffeine consumption and/or ED intake and self-reported behavioral executive function and metacognition are shown in **Tables 2A**,**B**. Consuming on average one ED or more a day was associated with the BRI (B 0.14, 95% CI 0.03; 0.24, $p = 0.012$), indicating more problems with self-reported behavior regulation. This association was due to the subscales measuring Inhibition and Monitor.

Participants who consumed at least one ED a day had higher scores on the MI (B 0.17, 95% CI 0.06; 0.29, $p = 0.003$). ED intake was associated with the metacognitive clinical subscales measuring Working Memory and Organization of Material. Participants who drank at least two consumptions of caffeine or ED had on average higher scores on the MI (B 0.09, 95% CI 0.01; 0.17, $p = 0.023$), indicating more problems with metacognitive skills. The effect estimate for this association was smaller compared to the effect estimate for the association between ED use and the MI. As the effect estimate of the sum of caffeine and ED use (B 0.09) falls within the CI of the association between ED use and the MI (95% CI 0.06;0.29), it is unlikely that there is a significant difference between the two associations. Furthermore, if there was any "cumulative effect," we expected the effect estimate of combined caffeine and ED use to be higher and not smaller than the one for the association between EDs and the MI. Finally, consuming at least one ED a day was related to more problems with behavior regulation and metacognition, whereas caffeine consumption was not related to any of the self-reported outcomes.

The adjusted associations between caffeine consumption and ED intake and parent reported behavioral executive function and metacognition are presented in **Tables 3A**,**B**. Only, consuming on average one or two caffeine containing drinks or EDs was associated with higher scores on the parent reported BRI (B 0.12, 95% CI 0.04; 0.20, $p = 0.005$), which was due to the association between caffeine and ED consumption and Inhibition. There were no statistically significant associations between caffeine use and/or EDs and parent reported metacognitive skills.

ADDITIONAL ANALYSES

Supplementary Tables 1 and 2 show how the magnitude of associations between caffeine and/or ED use and the indices changes by adding the potential confounders. "Models 1" show the unadjusted models. Adding gender, pubertal status and educational track, only slightly reduced the effect estimates.

DISCUSSION

The present study showed that during early adolescence consuming on average at least one ED a day was associated with more problems regarding behavior regulation and metacognition. Although caffeine use in the current sample was higher than ED consumption, we found no statistical significant associations between caffeine use and EFs. The sum of caffeine and ED consumptions was associated with self-reported problems with metacognition and with parent reported behavior regulation, but

Table 2A | Associations between caffeine and energy drink consumption and self-reported behavioral executive functioning ($n = 509$).

	Inhibition		Emotional control		Monitor		Shift		BRI[#]	
	***B* (95% CI)**	***p***	***B* (95% CI)**	***p***	***B* (95% CI)**	***p***	***B* (95% CI)**	***p***	***B* (95% CI)**	***p***
CAFFEINE										
<1 per day	Reference		Reference		Reference		Reference		Reference	
≥1–2 each day	0.06 (−0.02; 0.14)	0.13	−0.04 (−0.11; 0.04)	0.32	0.08 (−0.03; 0.18)	0.15	−0.03 (−0.10; 0.05)	0.50	0.02 (−0.05; 0.09)	0.58
≥2 each day	0.11 (0.01; 0.21)	0.05	0.01 (−0.08; 0.10)	0.75	0.04 (−0.08; 0.17)	0.49	0.01 (−0.09; 0.10)	0.94	0.04 (−0.04; 0.12)	0.33
EDs										
<1 per day	Reference		Reference		Reference		Reference		Reference	
≥1 each day	0.18 (0.05; 0.31)	0.008	0.07 (−0.05; 0.18)	0.28	0.22 (0.05; 0.38)	0.009	0.09 (−0.04; 0.21)	0.17	0.14 (0.03; 0.24)	0.012
CAFFEINE AND EDs										
<1 per day	Reference		Reference		Reference		Reference		Reference	
≥1–2 each day	0.09 (0.02; 0.17)	0.02	0.02 (−0.05; 0.08)	0.67	0.05 (−0.05; 0.14)	0.36	−0.02 (−0.09; 0.05)	0.59	0.03 (−0.03; 0.10)	0.30
≥2 each day	0.14 (0.05; 0.23)	0.05	0.02 (−0.06; 0.10)	0.66	0.08 (−0.03; 0.20)	0.15	0.01 (−0.08; 0.09)	0.90	0.06 (−0.01; 0.14)	0.10

Linear regression models adjusted for pubertal status, gender, educational track, number of sleeping problems and hours of sleep. Caffeine, EDs, and Caffeine and EDs represent the independent dummy-coded variables in which <1 consumption on average per day represents the reference category; self-reported behavioral executive functioning measures are the outcome variables. B, the estimate of increase in self-reported behavioral executive functioning score compared to the reference category; CI, Confidence Interval; [#]BRI, Behavior Regulation Index.

Table 2B | Associations between caffeine and energy drink consumption and self-reported metacognition ($n = 509$).

	Working memory		Planning and organizing		Organization of material		Persistence		MI[#]	
	***B* (95% CI)**	***p***	***B* (95% CI)**	***p***	***B* (95% CI)**	***p***	***B* (95% CI)**	***p***	***B* (95% CI)**	***p***
CAFFEINE										
<1 per day	Reference		Reference		Reference		Reference		Reference	
≥1–2 each day	0.01 (−0.08; 0.09)	0.88	−0.01 (−0.09; 0.07)	0.79	0.01 (−0.08; 0.11)	0.80	−0.02 (−0.11; 0.06)	0.58	0.00 (−0.08; 0.07)	0.91
≥2 each day	0.05 (−0.06; 0.15)	0.38	0.06 (−0.04; 0.15)	0.27	0.11 (−0.01; 0.23)	0.07	0.07 (−0.03; 0.18)	0.18	0.07 (−0.02; 0.16)	0.12
EDs										
<1 per day	Reference		Reference		Reference		Reference		Reference	
≥1 each day	0.19 (0.06; 0.33)	0.006	0.15 (0.03; 0.28)	0.017	0.20 (0.05; 0.35)	0.011	0.15 (0.01; 0.29)	0.031	0.17 (0.06; 0.29)	0.003
CAFFEINE AND EDs										
<1 per day	Reference		Reference		Reference		Reference		Reference	
≥1–2 each day	−0.01 (−0.09; 0.07)	0.82	0.00 (−0.08; 0.07)	0.99	0.04 (−0.05; 0.13)	0.39	−0.01 (−0.09; 0.07)	0.80	0.01 (−0.06; 0.07)	0.89
≥2 each day	0.10 (0.01; 0.19)	0.05	0.06 (−0.03; 0.15)	0.18	0.13 (0.03; 0.24)	0.015	0.09 (−0.01; 0.18)	0.08	0.09 (0.01; 0.17)	0.023

Linear regression models adjusted for pubertal status, gender, educational track, number of sleeping problems and hours of sleep. Caffeine, EDs, and caffeine and EDs represent the independent dummy-coded variables in which <1 consumption on average per day represents the reference category; self-reported metacognition measures are the outcome variables. B, the estimate of increase in self-reported metacognition score compared to the reference category; CI, Confidence Interval; [#]MI, Metacognition Index.

Table 3A | Associations between caffeine and energy drink consumption and parent reported behavioral executive functioning ($n = 319$).

	Inhibition		Emotional control		Shift		BRI−#	
	B (95% CI)	*p*	*B* (95% CI)	*p*	*B* (95% CI)	*p*	*B* (95% CI)	*p*
CAFFEINE								
<1 per day	Reference		Reference		Reference		Reference	
≥1–2 each day	0.11 (0.01; 0.21)	0.028	0.07 (−0.05; 0.18)	0.24	−0.02 (−0.12; 0.09)	0.74	0.05 (−0.04; 0.14)	0.24
≥2 each day	0.09 (−0.03; 0.20)	0.16	0.03 (−0.11; 0.17)	0.67	0.01 (−0.13; 0.14)	0.93	0.04 (−0.07; 0.15)	0.47
EDs								
<1 per day	Reference		Reference		Reference		Reference	
≥1 each day	0.04 (−0.13; 0.21)	0.64	0.00 (−0.21; 0.20)	0.96	−0.03 (−0.22; 0.16)	0.77	0.03 (−0.15; 0.16)	0.97
CAFFEINE AND EDs								
<1 per day	Reference		Reference		Reference		Reference	
≥1–2 each day	0.15 (0.06; 0.24)	0.002	0.13 (0.02; 0.24)	0.018	0.08 (−0.02; 0.18)	0.12	0.12 (0.04; 0.20)	0.005
≥2 each day	0.08 (−0.03; 0.18)	0.18	0.02 (−0.11; 0.15)	0.74	0.02 (−0.10; 0.14)	0.75	0.04 (−0.06; 0.14)	0.45

Linear regression models adjusted for pubertal status category, gender, educational track, number of sleeping problems and hours of sleep. Caffeine, EDs, and Caffeine and EDs represent the independent dummy-coded variables in which <1 consumption on average per day represents the reference category; parent reported behavioral executive functioning measures are the outcome variables. B, the estimate of increase in parent reported behavioral executive functioning score compared to the reference category; CI, Confidence Interval; #BRI, Behavior Regulation Index.

Table 3B | Associations between caffeine and energy drink consumption and parent-reported metacognition ($n = 319$).

	Initiative		Working memory		Planning and organizing		Organization of material		Monitor		MI #	
	B (95% CI)	*p*	*B* (95% CI)	*p*	*B* (95% CI)	*p*	*B* (95% CI)	*p*	*B* (95% CI)	*p*	*B* (95% CI)	*p*
CAFFEINE												
<1 per day	Reference		Reference		Reference		Reference		Reference		Reference	
≥1–2 each day	0.04 (−0.09; 0.15)	0.57	0.02 (−0.11; 0.15)	0.79	0.03 (−0.08; 0.14)	0.63	0.04 (−0.11; 0.20)	0.58	0.07 (−0.05; 0.19)	0.24	0.04 (−0.07; 0.14)	0.47
≥2 each day	0.12 (−0.03; 0.27)	0.11	0.09 (−0.07; 0.25)	0.26	0.12 (−0.02; 0.25)	0.09	0.19 (0.00; 0.38)	0.05	0.12 (−0.03; 0.27)	0.11	0.13 (0.00; 0.26)	0.05
EDs												
<1 per day	Reference		Reference		Reference		Reference		Reference		Reference	
≥1 each day	0.18 (−0.04; 0.39)	0.10	0.15 (−0.08; 0.37)	0.21	0.22 (0.03; 0.41)	0.026	0.14 (−0.13; 0.41)	0.30	0.22 (0.01; 0.43)	0.039	0.18 (0.00; 0.37)	0.05
CAFFEINE AND EDs												
<1 per day	Reference		Reference		Reference		Reference		Reference		Reference	
≥1–2 each day	0.02 (−0.10; 0.13)	0.78	0.04 (−0.08; 0.16)	0.52	0.07 (−0.03; 0.18)	0.17	0.08 (−0.07; 0.23)	0.28	0.15 (0.03; 0.26)	0.013	0.07 (−0.03; 0.17)	0.16
≥2 each day	0.12 (−0.02; 0.25)	0.09	0.08 (−0.07; 0.22)	31	0.12 (−0.01; 0.24)	0.07	0.12 (−0.05; 0.30)	0.16	0.10 (−0.04; 0.24)	0.14	0.11 (−0.01; 0.23)	0.08

Linear regression models adjusted for pubertal status category, gender, educational track, number of sleeping problems and hours of sleep. Caffeine, EDs, and Caffeine and EDs represent the independent dummy-coded variables in which <1 consumption on average per day represents the reference category; parent reported metacognition measures are the outcome variables. B, the estimate of increase in parent reported metacognition score compared to the reference category; CI, Confidence Interval; #MI, Metacognition Index.

the effect estimates of these associations did not seem to be statistically different from those of the associations between ED and EFs. Both effect estimates of combined use of caffeine and ED use fell within the confidence interval of the effect estimates of the association between ED use and EFs.

EDs, according to their manufacturers, enhance physical and, relevant to the current study, cognitive performance. Scientific studies investigating the effect of ED intake on cognitive performances either found improvement on cognitive tasks after consuming EDs (e.g., Lieberman, 2001; Wesnes et al., 2013) or found no evidence of an effect (e.g., Curry and Stasio, 2009; Wilhelm et al., 2013). In contrast, we found that ED intake was associated with an increased amount of problems with behavior regulation and metacognition. There are several potential explanations for the discrepancies in findings. First, fundamental differences in methods of assessment may have contributed to the differences between our findings and those of earlier studies. Previously conducted research focused on direct short-term effects of ED use on neurocognitive functioning assessed in experimental settings. ED use in this type of settings is by definition occasional use. We investigated the potential effect of habitual ED use on long-term cognitive functioning in daily life situations, measured by self-reports and parent reports. Although experiments enable researchers to precisely control the settings to measure certain cognitive functions, self-reports and parent reports give more insight in problems with executive functions in real life settings. Self-reports and parent reports, such as the BRIEF, measure executive function problems in a real-world setting and thus offer a higher ecological validity compared to laboratory based measures. In general, no or low correlations are reported between BRIEF measurements and performance-based measures (see for an overview in the literature: Huizinga and Smidts, 2011), which further illustrates that questionnaires and experimentally based measures are likely to tap different constructs (see also Toplak et al., 2013). Second, most previously conducted research studied the effects of EDs in older adolescents (Curry and Stasio, 2009; Wilhelm et al., 2013) or in samples with a broad age range (Wesnes et al., 2013). Our study focused specifically on young adolescents who are just entering their development toward adulthood, i.e., puberty. At the onset of puberty, the young adolescents' brain goes through a critical developmental phase, in which the maturation of the prefrontal brain areas plays a substantial role and has a large impact on a variety of cognitive functions. Our findings suggest that young adolescents who consume EDs on a regular basis perform worse in EFs than their non-using counterparts. Although the effects of caffeine consumption on brain development have not yet been examined, Temple (2009) hypothesizes that caffeine may alter normal brain development during critical developmental periods. This idea stems from animal models in which perinatal caffeine exposure had long lasting effects on brain function (For a review, see Temple, 2009). Third, it is possible that young adolescents that tend to consume caffeine and EDs may do so because of their - already - compromised EFs. EF may improve by caffeine and ED intake but may not be fully compensated in youngsters that use these drinks, i.e., in young adolescents who experience problems with EF. Finally, when interpreting the observed findings, it is important to take into account that 6% of our sample consumed on average at least one ED a day. Furthermore, the effect estimates of the associations between ED use and EFs were relatively small. These results may limit the impact of our findings.

Caffeine's stimulating effect on the central nervous system is well-established, and the capability of EDs to improve cognitive performances is often attributed to the caffeine they contain (Seifert et al., 2011). Therefore, we expected the associations between caffeine use and EFs vs. EDs and EFs to be similar. However, caffeine use was not associated with any of the outcome measures. Several explanations may underlie these findings. First, the quantity of caffeine in EDs can be substantially larger than in caffeine drinks. For example, the content of a regular cup of coffee usually varies between 125–250 ml, whereas cans of EDs vary between 250–500 ml. Second, in addition to caffeine, EDs contain high levels of sugar and smaller amounts of several other substances, such as vitamins, minerals, ginseng, taurine, inositol or other herbal extracts. In contrast to the idea that the effect of EDs is mainly due to the caffeine content, findings of other studies suggest that the combination of EDs' ingredients work synergistically (Scholey and Kennedy, 2004; Smit et al., 2004; Temple, 2009). These discrepancies in our findings, necessitate further research on EDs and ED components to determine their potential threats or benefits for health and performance.

Caffeine use and ED intake were moderately correlated. Therefore, we expected to find indications of a cumulative effect of combined use of caffeine and EDs. The effect estimate for the association between combined caffeine and ED use and parent reported BRI was, in absolute value, larger than the effect estimate for the association between EF and BRI, but fell within its confidence interval. The effect estimate for the association between combined caffeine and ED use and self-reported BRI was, in absolute value, even smaller than the effect estimate for the association between EDs and BRI, and also fell within its confidence interval. Therefore, it is implausible that the associations between the sum of caffeine consumptions and EDs are statistically different from the associations for caffeine and EDs separately. The moderate correlation between caffeine use and ED intake, indicating that few adolescents consumed both products, may have contributed to this finding.

STRENGTHS AND LIMITATIONS

This study has several strengths. First, this study is a first attempt to shed light on the potential long-term effects of ED intake on behavior or cognitive functioning in daily life situations, as previous research has focused on short-term direct effects of ED intake on cognitive functions. Second, in this study we investigated a large sample of specifically young adolescents entering or in puberty. We focused on this particular developmental phase as it can be determinative for later life functioning.

Several methodological limitations need to be discussed. First, participants reported the number of consumptions. Therefore, the exact amount of caffeine present in each consumption was unknown. Future research needs to focus on exact ED or caffeine intake by asking more detailed questions about the consumed brands and exact quantities of each consumption. Second, due to the cross sectional design of the study, we cannot ensure the

proposed directions of associations. It is possible that caffeine and EDs influence EFs, but the opposite is also feasible. As was mentioned earlier, young adolescents that use caffeine and EDs on a regular basis, may have certain characteristics, for example, compromised EFs beforehand, which may have influenced their ED intake instead of vice versa.

IMPLICATIONS AND CONCLUSIONS

Since the introduction of EDs, its market has grown immensely (Report Buyer, 2007). The prevalence of ED use has especially risen among children and adolescents (Reissig et al., 2009; Seifert et al., 2011; Zucconi et al., 2013). The fast rise in prevalence of ED use in this young population which is in a critical developmental phase is alarming, as there is a lack of research on the long term consequences of prolonged and regular use of these EDs on physical and cognitive health. Our findings suggest a possible negative effect of ED use on behavior regulation and metacognitive skills. However, it is important to note that in the current sample a relatively small group of young adolescents consumed on average EDs on a daily basis. Also, effect estimates were relatively small. We conducted a correlational study, therefore, no inferences can be made about causality or effects. Taken together, these results may limit the impact of our findings. Our findings do support the need for further more detailed research on the consequences of ED use in this vulnerable population. Future research on the possible negative effects of ED use on the more long-term EFs in daily life situations should focus on, for example on the exact amount of ED use, other sources of caffeine intake, and directions of associations.

AUTHOR CONTRIBUTIONS

Tamara Van Batenburg-Eddes performed the analyses for the current study, wrote the concept drafts and approved the final version of the manuscript, and takes full responsibility for all aspects of the work in terms of accuracy and integrity.

Nikki C. Lee had a substantial contribution to the interpretation of the data, writing and revising the manuscript, and approved the final version of the manuscript. She agrees to be accountable for all aspects of the work in terms of accuracy and integrity.

Wouter D. Weeda had a substantial contribution to the design of the work, analyses and interpretation of the data. Wouter D. Weeda contributed to revising the manuscript and approves of the final version to be published and agrees to be accountable for all aspects of the work in terms of accuracy and integrity.

Lydia Krabbendam is co-designer of the study, reviewed the manuscript and approved the final version of it and agrees to be accountable for all aspects of the work in terms of accuracy and integrity.

Mariette Huizinga is guarantor and designer of the study, supervised the first author, wrote and reviewed the manuscript and approved the final version of it, and takes full responsibility for all aspects of the work in terms of accuracy and integrity.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Initiative Brain and Cognition (NIHC 056-34-016, Mariette Huizinga). The authors gratefully acknowledge the contribution of participants, their parents and their schools. In addition, we thank Lisa van der Heijden, Daan Joosen, and Iris Dukevot for their substantial contribution to the data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00457/abstract

REFERENCES

Ames, S. L., Grenard, J. L., Thush, C., Sussman, S., Wiers, R. W., and Stacy, A. W. (2007). Comparison of indirect assessments of association as predictors of marijuana use among at-risk adolescents. *Exp. Clin. Psychopharmacol.* 15, 204–218. doi: 10.1037/1064-1297.15.2.218

Beebe, D. W., Fallone, G., Godiwala, N., Flanigan, M., Martin, D., Schaffner, L., et al. (2008). Feasibility and behavioral effects of an at-home multi-night sleep restriction protocol for adolescents. *J. Child Psychol. Psychiatry* 49, 915–923. doi: 10.1111/j.1469-7610.2008.01885.x

Bender, R., and Lange, S. (2001). Adjusting for multiple testing - when and how? *J. Clin. Epidemiol.* 54, 343–349. doi: 10.1016/S0895-4356(00)00314-0

Bramstedt, K. A. (2007). Caffeine use by children: the quest for enhancement. *Subst. Use Misuse* 42, 1237–1251. doi: 10.1080/10826080701208962

Carskadon, M. A., and Acebo, C. (1993). A self-administered rating-scale for pubertal development. *J. Adolesc. Health* 14, 190–195. doi: 10.1016/1054-139X(93)90004-9

Curry, K., and Stasio, M. J. (2009). The effects of energy drinks alone and with alcohol on neuropsychological functioning. *Hum. Psychopharmacol.* 24, 473–481. doi: 10.1002/hup.1045

Dahl, R. E., and Lewin, D. S. (2002). Pathways to adolescent health: sleep regulation and behavior. *J. Adolesc. Health* 31, 175–184. doi: 10.1016/S1054-139X(02)00506-2

Giedd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Ann. N. Y. Acad. Sci.* 1021, 77–85. doi: 10.1196/annals.1308.009

Gioia, G. A., Isquith, P. K., Guy, S. C., and Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function (BRIEF): Professional Manual.* Lutz, FL: Psychological Assessment Resources.

Gioia, G. A., Isquith, P. K., Retzlaff, P. D., and Espy, K. A. (2002). Confirmatory factor analysis of the behavior rating inventory of executive function (BRIEF) in a clinical sample. *Child Neuropsychol.* 8, 249–257. doi: 10.1076/chin.8.4.249.13513

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179. doi: 10.1073/pnas.0402680101

Graham, J. W., Flay, B. R., Johnson, C. A., Hansen, W. B., Grossman, L., and Sobel, J. L. (1984). Reliability of self-report measures of drug use in prevention research: evaluation of the Project SMART questionnaire via the test-retest reliability matrix. *J. Drug Educ.* 14, 175–193. doi: 10.2190/CYV0-7DPB-DJFA-EJ5U

Huizinga, M., Dolan, C. V., and Van Der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Huizinga, M., and Smidts, D. P. (2011). Age-related changes in executive function: a normative study with the dutch version of the behavior rating inventory of executive function (BRIEF). *Child Neuropsychol.* 17, 51–66. doi: 10.1080/09297049.2010.509715

Huizinga, M., and Smidts, D. P. (2012). *BRIEF Vragenlijst Executieve Functies Voor 5- tot 18-Jarigen.* Amsterdam: Hogrefe Uitgevers.

Jackson, M. L., Gunzelmann, G., Whitney, P., Hinson, J. M., Belenky, G., Rabat, A., et al. (2013). Deconstructing and reconstructing cognitive performance in sleep deprivation. *Sleep Med. Rev.* 17, 215–225. doi: 10.1016/j.smrv.2012.06.007

Lieberman, H. R. (2001). The effects of ginseng, ephedrine, and caffeine on cognitive performance, mood and energy. *Nutr. Rev.* 59, 91–102. doi: 10.1111/j.1753-4887.2001.tb06995.x

O'Brien, L. M. (2009). The neurocognitive effects of sleep disruption in children and adolescents. *Child Adolesc. Psychiatr. Clin. N. Am.* 18, 813–823. doi: 10.1016/j.chc.2009.04.008

Paus, T. (2013). How environment and genes shape the adolescent brain. *Horm. Behav.* 64, 195–202. doi: 10.1016/j.yhbeh.2013.04.004

Petersen, A. C., Crockett, L., Richards, M., and Boxer, A. (1988). A self-report measure of pubertal status - reliability, validity, and initial norms. *J. Youth Adolesc.* 17, 117–133. doi: 10.1007/BF01537962

Reissig, C. J., Strain, E. C., and Griffiths, R. R. (2009). Caffeinated energy drinks–a growing problem. *Drug Alcohol Depend.* 99, 1–10. doi: 10.1016/j.drugalcdep.2008.08.001

Report Buyer. (2007). *New Report Predicts Energy Drink Sales in the U.S. to Exceed $9 Billion by 2011*. Available online at: http://www.reportbuyer.com/press/new-report-predicts-energy-drink-sales-in-the-us-to-exceed-9-billion-by-2011/ (Accessed January 13 2014).

Roehrs, T., and Roth, T. (2008). Caffeine: sleep and daytime sleepiness. *Sleep Med. Rev.* 12, 153–162. doi: 10.1016/j.smrv.2007.07.004

Scholey, A. B., and Kennedy, D. O. (2004). Cognitive and physiological effects of an "energy drink": an evaluation of the whole drink and of glucose, caffeine and herbal flavouring fractions. *Psychopharmacology (Berl.)*. 176, 320–330. doi: 10.1007/s00213-004-1935-2

Seifert, S. M., Schaechter, J. L., Hershorin, E. R., and Lipshultz, S. E. (2011). Health effects of energy drinks on children, adolescents, and young adults. *Pediatrics* 127, 511–528. doi: 10.1542/peds.2009-3592

Smidts, D. P., and Huizinga, M. (2009). *BRIEF Executieve Functies Gedragsvragenlijst: Handleiding*. Amsterdam: Hogrefe.

Smit, H. J., Cotton, J. R., Hughes, S. C., and Rogers, P. J. (2004). Mood and cognitive performance effects of "energy" drink constituents: caffeine, glucose and carbonation. *Nutr. Neurosci.* 7, 127–139. doi: 10.1080/10284150400003041

Stichting Voedingscentrum Nederland (2013). *Energiedrankjes*. Available online at: http://www.voedingscentrum.nl/encyclopedie/energiedrankjes.aspx#blok10 (Accessed November 19, 2013).

Temple, J. L. (2009). Caffeine use in children: what we know, what we have left to learn, and why we should worry. *Neurosci. Biobehav. Rev.* 33, 793–806. doi: 10.1016/j.neubiorev.2009.01.001

Toplak, M. E., West, R. F., and Stanovich, K. E. (2013). Practitioner review: do performance-based measures and ratings of executive function assess the same construct? *J. Child Psychol. Psychiatry* 54, 131–143. doi: 10.1111/jcpp.12001

Wesnes, K. A., Barrett, M. L., and Udani, J. K. (2013). An evaluation of the cognitive and mood effects of an energy shot over a 6h period in volunteers: a randomized, double-blind, placebo controlled, cross-over study. *Appetite* 67, 105–113. doi: 10.1016/j.appet.2013.04.005

Wilhelm, P., Van Diepen, M. A., Nieuwenhuis, L., and Boulogne, T. L. (2013). The effect of energy drinks on the cognitive performance of adolescents. *Tijdschr. Psychiatr.* 55, 57–62.

Zucconi, S., Volpato, C., Adinolfi, F., Gandini, E., Loi, A., and Fioriti, L. (2013). *Gathering Consumption Data on Specific Consumer Groups of Energy Drinks*. European Food Safety Authority. Available online at: http://www.efsa.europa.eu/en/efsajournal/doc/394e.pdf

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2014; accepted: 29 April 2014; published online: 20 May 2014.
*Citation: Van Batenburg-Eddes T, Lee NC, Weeda WD, Krabbendam L and Huizinga M (2014) The potential adverse effect of energy drinks on executive functions in early adolescence. Front. Psychol. **5**:457. doi: 10.3389/fpsyg.2014.00457*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Van Batenburg-Eddes, Lee, Weeda, Krabbendam and Huizinga. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 17 June 2014
doi: 10.3389/fpsyg.2014.00593

Less-structured time in children's daily lives predicts self-directed executive functioning

Jane E. Barker[1]*, Andrei D. Semenov[1], Laura Michaelson[1], Lindsay S. Provan[1], Hannah R. Snyder[2] and Yuko Munakata[1]

[1] *Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA*
[2] *Department of Psychology, University of Denver, Denver, CO, USA*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Adele Diamond, The University of British Columbia, Canada
Angeline Lillard, University of Virginia, USA

****Correspondence:***
Jane E. Barker, Department of Psychology and Neuroscience, University of Colorado Boulder, 345 UCB, Boulder, CO 80309, USA
e-mail: jane.barker@colorado.edu

Executive functions (EFs) in childhood predict important life outcomes. Thus, there is great interest in attempts to improve EFs early in life. Many interventions are led by trained adults, including structured training activities in the lab, and less-structured activities implemented in schools. Such programs have yielded gains in children's externally-driven executive functioning, where they are instructed on what goal-directed actions to carry out and when. However, it is less clear how children's experiences relate to their development of self-directed executive functioning, where they must determine on their own what goal-directed actions to carry out and when. We hypothesized that time spent in less-structured activities would give children opportunities to practice self-directed executive functioning, and lead to benefits. To investigate this possibility, we collected information from parents about their 6–7 year-old children's daily, annual, and typical schedules. We categorized children's activities as "structured" or "less-structured" based on categorization schemes from prior studies on child leisure time use. We assessed children's self-directed executive functioning using a well-established verbal fluency task, in which children generate members of a category and can decide on their own when to switch from one subcategory to another. The more time that children spent in less-structured activities, the better their self-directed executive functioning. The opposite was true of structured activities, which predicted poorer self-directed executive functioning. These relationships were robust (holding across increasingly strict classifications of structured and less-structured time) and specific (time use did not predict externally-driven executive functioning). We discuss implications, caveats, and ways in which potential interpretations can be distinguished in future work, to advance an understanding of this fundamental aspect of growing up.

Keywords: cognitive development, self-directed executive function, leisure time, unstructured activities, verbal fluency

INTRODUCTION

Why do young children often forget (or outright refuse) to put on a coat before leaving the house on a snowy day? The choice to put on a jacket may seem frustratingly obvious to parents and older siblings, but this simple decision arises from a surprisingly complex interplay of behaviors. Children must keep in mind a goal (staying warm and dry) that is not yet relevant in the comfort of a warm house. They must inhibit the urge to proceed with a regular sequence of tasks (put on socks and shoes and head out the door), and instead modify their routine to include something new (pulling a coat from the closet). Unless someone intervenes, this change in the status quo must be accomplished without any external reminders (a visible coat, or a well-timed reminder from a caregiver).

To accomplish each of these tasks, children must engage executive functions (EFs), the cognitive control processes that regulate thought and action in support of goal-directed behavior. EFs develop dramatically during childhood (e.g., Gathercole et al., 2004; Zelazo et al., 2008; McAuley et al., 2011; Munakata et al., 2012), and support a number of higher-level cognitive processes, including planning and decision-making, maintenance and manipulation of information in memory, inhibition of unwanted thoughts, feelings, and actions, and flexible shifting from one task to another. Researchers have used a variety of laboratory tasks to measure child EFs, including tabletop behavioral tasks (e.g., the classic marshmallow test, card-sorting tasks) and computerized tasks (e.g., Go/No-go, Flanker), many of which tap multiple aspects of EF. Over the past decade, EFs have emerged as critical, early predictors of success across a range of important outcomes, including school readiness in preschoolers (Miller et al., 2013), as well as academic performance at school entry (Blair and Razza, 2007; Cameron et al., 2012) and beyond (St Clair-Thompson and Gathercole, 2006; Best et al., 2011). Moreover, children with worse EF go on to have poorer health, wealth, and social outcomes in adulthood than children with better EF, even after

controlling for differences in general intelligence (Moffitt et al., 2011).

Given the established links between early EFs and later life outcomes, a number of studies have investigated whether EF abilities can be changed through experience, with some notable successes. Most of this work has involved adult-led training or interventions, which allow children to practice EFs in an environment where adults provide some guidance. For example, children's working memory, or their ability to maintain and manipulate information across a delay, can be improved through short periods of targeted training (e.g., Holmes et al., 2009; Bergman Nutley et al., 2011). During such training, children are presented with sequences of spoken or visual stimuli. After a brief pause, the child is instructed to reproduce the sequence either in forward order (requiring maintenance of information, but no manipulation) or in reverse order (requiring maintenance and manipulation). After training, children often show better performance in similar tasks assessing the same skills (e.g., Holmes et al., 2009, 2010; Thorell et al., 2009; Bergman Nutley et al., 2011; reviewed in Diamond and Lee, 2011; Shipstead et al., 2012). In addition, children's cognitive flexibility, or ability to change tasks or strategies in response to new environmental demands, can be improved via interventions implemented in preschool curricula (Lillard and Else-quest, 2006; Diamond et al., 2007; Bierman et al., 2008; Röthlisberger et al., 2012). These curricula have ranged from partial-day, small-group sessions where children play games developed to exercise EFs (Röthlisberger et al., 2012), to comprehensive, full-day implementations, such as those found in Tools of the Mind (Diamond et al., 2007) and Montessori (Lillard and Else-quest, 2006) classrooms, where teachers are trained to scaffold developing EFs throughout the day. Relative to children in business-as-usual classrooms, children enrolled in such curricula have subsequently shown better performance in tasks where they must flexibly shift from one rule (e.g., sorting cards by their shape) to another (e.g., switching to sorting the cards by color).

Altering children's experiences with such training and interventions has thus led to improvements in children's externally-driven EF, where they are instructed on what to do (e.g., sort cards according to shape; remember a sequence of digits), and when (e.g., now switch and sort according to color; now recall the digits in reverse order). In the real world, children who have developed externally-driven EF might behave in a goal-directed way when given reminders. For example, a child might successfully put on a coat in the morning after a reminder from a caregiver. However, it is less clear how children's experiences relate to their development of more self-directed executive functioning, where they must determine on their own what goal-directed actions to carry out and when. A self-directed child, for example, might put a coat on just before going outside without being told what to do.

The development of self-directed EF is a critical part of growing up. Self-directed EFs develop later than externally-driven forms of executive control (Welsh et al., 1991; Jacques and Zelazo, 2001; Smidts et al., 2004; Snyder and Munakata, 2010; Chevalier et al., 2011), and prove to be more cognitively demanding, even in adults (e.g., Bryck and Mayr, 2005; Forstmann et al., 2005; Lie et al., 2006). Tasks assessing self-directed control typically provide an overall goal, but challenge participants to generate their own rules for how and when to employ EFs to achieve that goal. For example, in the verbal fluency task, which is a frequently-used and longstanding measure of EF (e.g., Troyer et al., 1997, 1998; Henry and Crawford, 2004; Sauzéon et al., 2004; Costafreda et al., 2006; Birn et al., 2010; Raboutet et al., 2010; Unsworth et al., 2010), participants are given a category (e.g., foods), and asked to produce as many words falling within that category as possible across a 1-min interval. To produce many items, participants may cluster responses (by grouping words that fall within the same semantic subcategory) and switch between subcategories when available exemplars are in short supply (e.g., Troyer et al., 1997, 1998; Abwender et al., 2001; Koren et al., 2005). Individuals must endogenously detect the need to switch (e.g., when they cannot think of more breakfast foods) and select what to switch to (e.g., desserts, vegetables, or fruits). Each process critically relies on generation of internal cues and becomes less executively demanding when external cues are instead provided (Randolph et al., 1993; Tremblay and Gracco, 2006; Snyder and Munakata, 2010). Consistent with this analysis of the self-directed nature of this task, switching among subcategories has been well-validated as the most executively-demanding component of verbal fluency tasks: switching (as opposed to naming items within clusters) activates prefrontal cortex (e.g., Hirshorn and Thompson-Schill, 2006), is impaired by prefrontal lesions (e.g., Troyer et al., 1998), and has the most protracted developmental course, with performance continuing to increase through adolescence (e.g., Kave et al., 2008). Young children often fail to switch from one subcategory to another, and instead perseverate on an initial subcategory (e.g., naming five different breakfast foods, and then indicating to the experimenter that they are finished). However, like patients with frontal lobe dysfunction, who benefit from semantic cueing during verbal fluency (Randolph et al., 1993; Drane et al., 2006; Iudicello et al., 2012), children can improve on the task when demands on self-directed EF are reduced by providing example subcategories prior to the task (Snyder and Munakata, 2010). This body of literature highlights the role of self-directed EF in switching among subcategories in the verbal fluency task.

We predicted that children's self-directed EFs might benefit from participation in less structured activities, where children, rather than adults, choose what they will do and when. Such experiences could support the practice of self-directed executive functioning, and lead to benefits. For example, children may practice engaging self-directed forms of EF by establishing goals and carrying them out across an afternoon ("first I'll read this book, then I'll make a drawing about the book, then I'll show everyone my drawing") or during a visit to a museum ("first I want to see the dinosaur exhibit, and then I want to learn about rocks"). These types of self-directed choice and planning are central to the Tools of the Mind and Montessori classrooms, although the exact form they take and the types of activities emphasized differ across these programs (Montessori, 1976; Bodrova, 2003; Bodrova and Leong, 2007).

For example, extended, social pretend play figures centrally in the Tools of the Mind program. This program is based on the work of Vygotsky (Bodrova and Leong, 2007), who theorized that imaginative play supports the development of self-directed EF, in children's transitions from other-regulated to self-regulated

cognitive processes (Vygotsky, 1967). During pretend play, children may practice engaging self-directed forms of EF by developing and maintaining their own goals to guide their behavior, even in the presence of conflicting environmental signals: a child who uses a wooden spoon as a wand maintains a pretend use while inhibiting a typical use (stirring a pot). Children's pretend play, as assessed during laboratory tasks with an adult experimenter, does predict their externally-driven EFs (Albertson and Shore, 2009; Kelly et al., 2011; Carlson et al., 2014); however, this relationship has been observed less reliably when pretend play is assessed during naturalistic play (e.g., Elias and Berk, 2002; Kelly et al., 2011; cf. Harris and Berk, as discussed in Lillard et al., 2013).

While preschool programs such as Tools of the Mind and Montessori implement the types of activities that we predict will benefit self-directed EFs, and such programs improve children's externally-driven EFs as discussed above, little work has investigated the relationship between such activities and the development of self-directed EFs. One study found that 12-year-old Montessori students were rated more highly on a measure of creativity than non-Montessori students, when writing answers to complete the prompt, "__ had the best/worst day at school" (Lillard and Else-quest, 2006). While such findings are suggestive because open-ended writing assignments have the potential to tap self-directed EFs, the prompt completion task is not an established measure of self-directed EFs, and there is some debate about the extent to which creativity reflects EF (e.g., Groborz and Nęcka, 2003; Chrysikou and Thompson-Schill, 2011; Ellamil et al., 2012; Jarosz et al., 2012). In addition, improved performance on this task may also reflect other benefits from such programs to language or writing skills; additional benefits were in fact observed on this task in the Montessori students' sentence sophistication. Moreover, it is unclear whether a broader range of less-structured activities outside of formal schooling yield EF benefits. Investigating this question is important, given that effects observed inside formal settings with trained adults may not generalize to other settings (as in the case of the pretend play effects discussed above), and given that not all families have access to the school settings where effects have been observed.

As a first step in examining the question of how children's experiences outside of formal schooling relate to EFs, we conducted a naturalistic, correlational study, in which we measured the time that 6-year-old children spent in their daily lives in structured and less-structured activities and tested whether it predicted performance in the lab on well-established executive function tasks, both externally-driven and self-directed. At this age, children spend some time in both structured and less-structured activities (e.g., Meeks and Mauldin, 1990; Hofferth and Sandberg, 2001a) and show some ability in self-directed control tasks, without showing high levels of proficiency (e.g., Welsh et al., 1991; Brocki and Bohlin, 2004; Kave et al., 2008; Snyder and Munakata, 2010, 2013).

To classify structured and less-structured activities, we relied on studies of child leisure time use (e.g., Meeks and Mauldin, 1990; Larson and Verma, 1999; Hofferth and Sandberg, 2001b; Fletcher et al., 2003; Osgood et al., 2005), which have attempted to discriminate between activities constituting structured, or constructive leisure, and "unstructured" leisure activities. "Unstructured" activities in this literature might be better thought of as "less-structured" activities, given that they can include some adult structuring, so we use the latter terminology throughout this paper. Most leisure time studies have identified structured leisure activities as those that are "supervised to some degree by a conventional adult, are highly structured, and provide [children] with a clear set of conventional activities in which to engage" (Agnew and Petersen, 1989, p. 335). Such activities "are…organized by adults around specific social or behavioral goals" (Fletcher et al., 2003, p. 641). Thus, structured time in the present study was defined to include any time outside of formal schooling[1] spent in activities organized and supervised by adults (e.g., piano lessons, organized soccer practice, community service, homework). Less-structured activities have been described more loosely, and generally include voluntary leisure activities where adults provide fewer guidelines or direct instructions, such as activities that are "spontaneous, [taking] place without formal rules or direction from adult leaders, and [featuring] few goals related to skill development" (Mahoney and Stattin, 2000, p. 116). Our coding scheme follows existing coding schemes documented in Meeks and Mauldin (1990) and Hofferth and Sandberg (2001b). In cases where these coding schemes differed, we reviewed the literature to ensure that our coding was in accordance with the majority of other time use studies[2]. In the present study, less-structured activities included activities such as free play, family and social events, reading, drawing, and media time. While these classifications are imperfect (e.g., they do not capture the degree of structure within and across classifications—an issue we return to in the Discussion), they allow us to build on the existing literature, and serve as an important starting point for testing our predictions; further analyses allow us to test the importance of particular activities within these classifications.

We hypothesized that the amount of time children spent in less-structured activities would predict their self-directed EF, over and above any differences attributable to age, general vocabulary knowledge, and household income. We expected these effects to be specific, such that less-structured activities would not predict externally-driven EF and structured activities would not predict self-directed EF.

METHODS

PARTICIPANTS

Seventy children participated in the study [M_{age} = 6.58 years; range = (6.01–7.00 years); males = 37]. All participants were

[1]We did not classify time spent in school as "structured" because the degree of structure in school settings can vary a great deal, and parent reports are likely to be inaccurate (since parents often do not have direct knowledge of child activities during schooling hours). Our definition of structured activities is also consistent with past studies of structured leisure time, which have excluded time spent in school (e.g., Meeks and Mauldin, 1990; Larson and Verma, 1999; Mahoney and Stattin, 2000; Hofferth and Sandberg, 2001b; Fletcher et al., 2003; Osgood et al., 2005).

[2]Hofferth and Sandberg (2001b) separately identify reading, studying, and television watching as learning activities. However, we have classified reading and television as less-structured time, and studying as structured time, in keeping with other studies (e.g., Meeks and Mauldin, 1990; Eccles and Barber, 1999; Fletcher et al., 2003).

recruited from a database of families who had volunteered to participate in research. During subject recruitment, parents were informed that they would be asked to document child activities during the week prior to the study visit. Three participants were excluded from analyses because detailed information on their weekly activities was unavailable, either because parents did not wish to provide this information (2), or because data were lost due to a technical error at the time of parent submission (1). Of the remaining participants, one child did not complete the Flanker task, one child did not complete the digit span task, and two children did not complete the verbal fluency task; each of these children was excluded from the analysis of only that task. All other participants completed all study tasks. Prior to their participation, parents gave informed consent, and children gave verbal assent. Children received small gifts (e.g., gliders, balls) throughout the project for their participation, and parents received $5 as compensation for travel.

DESIGN AND PROCEDURE

Children were individually tested in a single session lasting approximately 1.5 h, with breaks given as needed. All children completed tasks in the same order: AX-CPT, Flanker, forward digit span (for other purposes, not discussed further in this report[3]), verbal fluency, and the Expressive Vocabulary Test. During the child tasks, parents provided demographic information and completed surveys of children's daily, annual, and typical schedules, as well as an exploratory "helicopter parenting" scale (not discussed further in this report; from Obradovic, pers. commun., October 26, 2011).

Parent questionnaires

Parent survey of child time use. Parents reported all child activities during the week prior to the laboratory test session using a computer-based survey. At the time that the study visit was scheduled, parents were informed that they would complete a detailed child activity survey during their visit, and were encouraged to take notes on their child's activities throughout the week. Parents were allowed to consult notes as they completed the survey. The survey was formatted as a 36 × 7 grid, such that each cell represented a 30-min time interval during the prior week (intervals occurring between midnight and 5:30 am were excluded to reduce burden). In each cell, parents wrote short, open-ended description of their child's activities, excluding times where children were sleeping or in school (parents indicated sleep and school schedules in a separate section of the survey). Before completing the survey, parents were asked to indicate the extent to which their family's activities over the prior week reflected typical patterns of time use. Parents rated their level of agreement with the prompt, "Was your family's schedule last week unusual or atypical?" via a 7-point scale anchored by "Strongly agree" and "Strongly disagree." Parents were then given verbal and written instructions, as follows:

> "Be as specific as possible for every activity you report. For example, for time spent in the car during a commute, rather than writing, "Drove from ___ to ___," you could write, "Watched a DVD with his sister in the car while driving to the city for a research appointment."
>
> Indicate who your child was interacting with during a given activity. For example, if your child had free time to play outside between dinner and bedtime, rather than writing "Free time outside," you could write, "Played tag outside with older sister and friends from the neighborhood." Or, if your child reads before bedtime, rather than writing, "Reading time," you could write, "Read aloud to mom before bed."
>
> Indicate simultaneous activities. For example, if your child ate a snack after school or camp while he/she had some down time, rather than writing "Snack time," you could write, "Ate a snack while coloring."

As parents completed the survey, experimenters periodically reviewed responses and asked that parents modify entries that were difficult to interpret or insufficiently detailed. Experimenters were also available during breaks between tasks to respond to parent questions about specific responses.

Child activity data were coded by three independent raters who were blind to data on all other tasks during each stage of the coding process. Coders assigned a numeric code to each cell-based survey entry using an activity classification scheme (**Table 1**). To ensure consistency across raters and reduce procedural drift, all raters independently classified each cell for the first 35 participants. Coders then met to discuss major discrepancies and to generate additional generalizable rules. Coders categorized responses from the final 32 participants using these agreed-upon criteria. The final 32 subjects were used to establish inter-rater reliability; reliabilities among pairs of coders ranged from 0.96 to 0.97, with coders agreeing on 7942 to 8021 cells out of 8288 total (i.e., 2 cells per hour × 18.5 h/day × 7 days a week × 32 participants). Excluding sleep and school cells (where there were no discrepancies between coders), reliabilities among pairs of coders were also high, ranging from 0.93 to 0.95. The three coders met to discuss discrepancies and generate a final, coded data set for each participant.

After the raters generated the final set of activity codes, each activity was further classified as either "Structured" or "Less-Structured" based on the coding scheme outlined in **Table 1**, following existing coding schemes (Meeks and Mauldin, 1990; Eccles and Barber, 1999; Mahoney and Stattin, 2000; Hofferth and Sandberg, 2001b; Fletcher et al., 2003; Osgood et al., 2005). All child-initiated activities (play, spontaneous practice, reading, watching television) and outings and events (museum or library visits, sporting events) were coded as "Less-Structured." Adult-led lessons and practices, homework and studying, religious activities, and organization meetings (e.g., community service) were coded as "Structured."

Parent survey of typical child time spent in less-structured activities. In a separate survey, parents were asked to indicate how often their children engaged in typical play activities by using a 7-point scale ("Never," "Less than once a month," Once a month," "2–3 times a month," "Once a week," "2–3 times a week," "Daily") to rate the following items: "Surf the

[3]Forward digit span tasks (where children repeat numbers in the order they are presented by an experimenter) primarily index storage capacity, rather than combined storage and processing capacity, and therefore do not serve as a reliable measure of EF (Daneman and Merikle, 1996; Engle et al., 1999).

Table 1 | Classification of child time use (structured, less-structured, and other activities).

STRUCTURED ACTIVITIES
Physical lessons (e.g., soccer practice, karate)
Non-physical lessons (e.g., piano lessons, art class)
Tutoring
Homework and study
Chores
Religious activities
Other formal organizational meetings and activities (e.g., community service)
LESS-STRUCTURED ACTIVITIES
Unguided, child-initiated practice (e.g., playing piano or singing outside of scheduled practice times; shooting goals outside of soccer practice)
Free play alone
Free play with others
Social outings
Visits to family and friends
Parties
Camping
Picnics
Other group activities (e.g., walks, bike rides, skiing, swimming, bowling, golf)
Enrichment activities
Sightseeing
Aquarium and zoo visits
Museums
Miscellaneous educational events (e.g., science fair)
Other entertainment (e.g., live sporting events, performances, movies)
Reading
Media and screen time (e.g., TV, internet, video games)
OTHER ACTIVITIES
Sleeping
Meals/eating
School
Care by others
Personal care and hygiene
Child appointments
Commuting and travel time
Unknown/Unreported

All entries that parents provided in the child time use survey were classified into these categories, following existing coding schemes (Meeks and Mauldin, 1990; Eccles and Barber, 1999; Mahoney and Stattin, 2000; Hofferth and Sandberg, 2001b; Fletcher et al., 2003; Osgood et al., 2005).

internet," "Watch television, videos/DVD, or online media," "Play video games (non-instructional)," "Play interactive instructional or learning games," "Play with toys alone," "Play with toys with friends/siblings," "Play physical games with friends/siblings," "Play physical games alone," "Play non-physical games alone," "Play card or board games with family," "Read," "Help with housework or cooking," "Play musical instrument", "Listen to music." Scores on each item (where 1 = "Never" and 7 = "Daily") were summed to produce a typical less-structured activity score.

Parent survey of seasonal child activities. In a separate survey, parents were asked to indicate the number of hours their child spent in structured lessons during the past year. Parents responded to 18 common structured lessons (basketball, baseball, tennis, hockey, soccer, football, golf, swimming, dance, gymnastics, martial arts, skiing/snowboarding, ice skating, music, art, theater, and tutoring) and were asked to write in any structured lessons that did not fall into these categories (most commonly, religious activities, and organizational meetings). To reduce burden, parents provided seasonal time estimations for each activity (e.g., the typical hours per week a child spent participating in music lessons over the prior fall). Data were reviewed for accuracy to ensure that parent-reported structured activities adhered to the same coding guidelines used to evaluate the Parent Survey of Weekly Activities. Cumulative hours spent in structured activities across the year were summed to produce an annual structured hours score.

Household income. Parents reported annual household income via an interval scale (median bracket: $100,000—$124,999; range: < $25,000 to > $150,000 USD). Fourteen parents chose not to disclose income information.

Child endogenous executive function measure

Verbal fluency. In the verbal fluency task, children were asked to generate words in response to a categorical prompt. The task was presented as a game to make it more engaging for children (as in Snyder and Munakata, 2013). Children were told, "We're going to play a game where we think of lots and lots of words. I bet you're really good at thinking of words, aren't you? I'll tell you what kinds of words to think of, and every time you tell me one, I'll put a pom-pom in your cup. Let's see how many pom-poms you can get before all the sand is gone (experimenter pointed to a 1-min sand timer children could use to estimate how much time was left). I'll bet you can get a lot! And when we are all done thinking of words, you can trade the pom-poms for a prize." Before each category, the experimenter said, "This time I want you to tell me as many [category name] as you can think of. Can you think of lots and lots of [category name]? Ready, go!" The experimenter placed a pom-pom in a clear plastic cup in front of the child for each new exemplar. If children paused for 10 s or longer between items, they were encouraged to continue ("Good job, can you tell me some more [category name]?"). In the rare instance where a child stated that she/he had named all words, the experimenter double-checked with the child (e.g., "Are you sure? What other [category name] can you think of?") and waited with the child until the end of the block. Children completed three blocks using this procedure, each of 1-min duration: a practice block (with the prompt "household items"), and two test blocks (with the prompts "animals" and "foods," which were counterbalanced across participants).

Verbal fluency data were transcribed from audio recordings, and coded by the experimenter and two independent raters blind to data on all other tasks. Coders identified clusters of items that were semantically related (e.g., "cookies, pie, cake" when producing foods). Switches between clusters of related items were identified and summed to generate cumulative switch scores. Switch scores were weighted by cluster size (as in Snyder and Munakata, 2010, 2013), such that 1 point was awarded for a switch after a cluster of 2 related items, 2 points for a switch after 3 related items, 3 points for switch after 4 related items, and so on. Weighted switch scores were used because they reflect increasing confidence as cluster size increases that children are indeed clustering and

switching. Unweighted scoring systems (e.g., Troyer et al., 1997), which count every transition between subcategories equally (including between single, unclustered items), have been criticized for confounding switching with a failure to cluster (e.g., Abwender et al., 2001). Inter-rater reliabilities were high between all pairs (>85%). To generate cumulative switch scores for each participant, weighted switch scores were averaged across coders within each prompt, and then summed.

Child externally-driven executive function measures

Flanker. Children completed a computerized flanker task (Eriksen and Schultz, 1979) assessing their ability to resolve conflicting visual information by appropriately responding to a central stimulus while ignoring flanking stimuli. The Flanker task is a commonly-used measure of externally-directed EF in 6-year-olds (e.g., Ridderinkhof and van der Molen, 1995; Rueda et al., 2004, 2005; McDermott et al., 2007; Röthlisberger et al., 2012) and has been shown to be sensitive to some interventions targeting EF in this age group (Fisher et al., 2011; Röthlisberger et al., 2012). During the task, children were instructed to indicate the orientation (left or right pointing) of a centrally-presented target stimulus, via a corresponding button press. In congruent trials, the target stimulus (the center fish) was surrounded by fish with the same orientation. In incongruent trials, the target image was surrounded by fish with an opposite orientation. In neutral trials, only the target image was presented and was not surrounded by any fish. Following a 10-trial practice block (4 congruent, 4 incongruent, 2 neutral), children completed three 32-trial blocks of the task: two incongruent blocks (for each block, incongruent trial $N = 16$; neutral trial $N = 16$), separated by one congruent block (congruent trial $N = 16$; neutral trial $N = 16$). Trials were presented in random order within blocks.

Reaction times were used to assess children's ability to resolve interference among conflicting stimuli, as in past work with this age group (e.g., Rueda et al., 2005; McDermott et al., 2007; Röthlisberger et al., 2012). Incongruent trials require children to attend to only the target middle fish and to ignore the surrounding fish. Therefore, the flanker task can be used to assess children's ability to filter out irrelevant information. Larger interference costs (i.e., the difference between average response time on incongruent trials and average response time on neutral trials) reflect greater difficulty filtering irrelevant information. To assess filtering ability, we first calculated participant mean response times for each trial type (neutral, incongruent/congruent) within each block across trimmed, correct trials (trials <100 and >3000 ms were excluded, as well as any trials three standard deviations outside that participant's mean for that trial type and block). To generate robust estimates of possible interference effects (as suggested by Lavie, 1995, and implemented in D'Ostilio and Garraux, 2012), incongruent/congruent trial mean RTs were contrasted with neutral trial mean RTs from the same block, yielding one congruent-neutral contrast and two incongruent-neutral contrasts within each participant. Flanker conflict scores were generated by subtracting the congruent contrast from each incongruent contrast (yielding two conflict scores, one arising from each incongruent block). These conflict scores were averaged to generate a summary flanker conflict score.

AX-CPT. Children completed the AX Continuous Performance Task (AX-CPT), which provides a measure of proactive control, or the tendency to maintain goal-relevant information until it is needed (Braver et al., 2007). All procedures and analyses were conducted as in Chatham et al. (2009). In this touchscreen-based, child-friendly version, children are allowed to prepare for future circumstances (the appearance of either "X" or "Y" image probes) based on previous experiences (the appearance of "A" or "B" image cues).

Children were instructed to respond with a target response whenever the "A" context cue was followed by an "X" probe. Children were instructed to provide a non-target response to all other cue-probe sequences (A – Y; B – X; B – Y). To improve child engagement during the task, popular cartoon characters were used as image stimuli, and the instructions took the form of character preferences. For example, children were told, "Spongebob likes watermelon, so press the happy face when you see Spongebob and then the watermelon," and, "Blue doesn't like the slinky, so press the sad face when you see Blue and then the slinky."

After the experimenter explained the task rules, children completed a "verification" phase to ensure that they understood the instructions and were capable of following rules. During this phase, each cue–probe pair was presented sequentially, and participants were asked to indicate the correct response for each pair. If subjects responded incorrectly to a cue-probe pair, the experimenter repeated the relevant rule ("Remember, when you see [A, B] and then you see [X, Y], tap this button [appropriate button blinks] as quickly as you can!") and subjects were allowed to try again. Participants then completed 7 practice trials. Cues were presented for 500 ms, followed by a 120 ms delay period, and a subsequent 6 s probe, as in test trials. Test trials were presented in four 30-trial blocks, where 70% of trials were target (A – X) trials, and 30% were non-target trials (A – Y; B – X; B – Y, appearing in equal proportion).

Proactive children show a characteristic behavioral profile that can be used to generate an RT-based measure of proactive control. Children who engage proactive control generate fast RTs in BX and BY trials, since maintenance of the "B" cue supports a non-target response to the subsequent "X" probe, and slower RTs on AY trials, since active maintenance of the "A" cue leads to anticipation of an "X" probe (due to the expectancy generated by asymmetric trial type frequencies). Proactive control was thus calculated using the median of trimmed RTs on correct AY and BX trials, which were entered into the formula (AY – BX)/(AY + BX). All responses made <200 ms after the presentation of the probe were removed from the analysis, resulting in the exclusion of <1% of all trials.

Expressive vocabulary test. The EVT (Pearson Assessments, Bloomington, MN) is a standardized, nationally normed, expressive vocabulary test, which we used (as in Snyder and Munakata, 2010) to control for differences in vocabulary that might have influenced verbal fluency performance (i.e., a child with a robust vocabulary might be capable of generating larger clusters than a child with a limited vocabulary, independent of either child's switching ability). On each trial of the EVT, children are shown a colored picture and are asked to name it or provide

a synonym (e.g., "Can you tell me another word for father?"). Testing continues until children incorrectly answer five items in a row, and raw scores are then converted into a standardized score based on age.

RESULTS

PRELIMINARY RESULTS AND ANALYSIS APPROACH

Weekly and annual/typical estimates of how children spent their time (**Figures 1A–C**) were marginally correlated, for both structured activities ($r = 0.24$; $p < 0.06$) and less-structured activities ($r = 0.23$; $p < 0.071$). We thus generated composite scores across weekly and annual/typical estimates to provide a more accurate and reliable measure of children's time. Each composite measure (for structured time, and separately for less-structured time) was formed by summing z-scored time in prior-week activities with z-scored ratings from the parent survey of annual/typical child activities, within each participant.

All analyses were conducted using standard linear regression. We included age, gender, and family income as factors in all models, given that they or related factors are often predictive of children's EF: age (e.g., Welsh et al., 1991; Huizinga et al., 2006), gender (e.g., Blair et al., 2005; Diamond et al., 2007), family income Hughes et al., 2009; as a component of SES: (Farah et al., 2006; Noble et al., 2005, 2007; Raver et al., 2013). Child vocabulary, as indexed by EVT performance, was included as a covariate in all tests of verbal fluency performance. Descriptive statistics for executive function, vocabulary, and time use measures are given in **Table 2**. Individual EF measures were not correlated, before or after controlling for age (p's > 0.4). For all analyses, outlying observations were identified (Cook's D > 3 standard deviations above the mean) and removed. This resulted in the exclusion of no more than four cases from any analysis.

CHILD TIME USE AND SELF-DIRECTED EF

Less-structured time

As predicted, children who spent more time in less-structured activities demonstrated better self-directed EF, as indexed by verbal fluency performance [$\eta_p^2 = 0.07$; $F_{(1,\ 44)} = 4.46$; $p < 0.05$; **Figure 2A**; **Table 3**]. In addition, older children and children with higher vocabulary scores demonstrated better verbal fluency performance [Age: $\eta_p^2 = 0.11$; $F_{(1,\ 44)} = 7.45$; $p < 0.01$; EVT: $\eta_p^2 = 0.10$; $F_{(1,\ 44)} = 6.30$; $p < 0.02$]. In subsequent tests for interactions, we found an unexpected interaction between less-structured time and age [Less-structured time × Age: $\eta_p^2 = 0.08$; $F_{(1,\ 43)} = 5.48$; $p < 0.03$]. *Post-hoc* tests indicated that additional time in less-structured activities predicted better self-directed control in most but not all children; specifically, this finding held in both the youngest sample quartile [$M_{age} = 6.38$ years, Less-structured time: $\eta_p^2 = 0.07$; $F_{(1,\ 43)} = 10.37$; $p < 0.003$] and at the median [$M_{age} = 6.65$ years, Less-structured time: $\eta_p^2 = 0.07$; $F_{(1,\ 43)} = 6.81$; $p < 0.02$], but not in the oldest quartile ($M_{age} = 6.86$ years; $p > 0.8$). When the interaction between less-structured time and age was included in the model, children from higher-income households demonstrated marginally better verbal fluency performance {Income: $\eta_p^2 = 0.05$; [$F_{(1,\ 43)} = 3.36$; $p < 0.08$]}. Age, vocabulary, and time in less-structured activities also continued to predict self-directed EF [Age: $\eta_p^2 = 0.12$; $F_{(1,\ 43)} = 5.76$; $p < 0.03$; Vocabulary: $\eta_p^2 = 0.07$; $F_{(1,\ 43)} = 4.80$; $p < 0.04$; Less-structured time: $\eta_p^2 = 0.07$; $F_{(1,\ 43)} = 6.81$; $p < 0.02$].

Exploratory analyses. We next investigated whether specific kinds of less-structured activities were driving the observed relationship between less-structured time and self-directed control. Composite variables representing common less-structured activities were created by aggregating similar responses across prior-week and annual/typical measures[4]. This procedure yielded seven broad categories of less-structured activities: unguided practice; play alone; play with others; social events with family (including parties, camping, picnics, and other group outings, such as hiking, biking, and swimming[5]), enrichment events (visits to the museum, library, aquarium, or zoo; sightseeing; and miscellaneous educational events), other entertainment (movies, performances, and live sporting events); reading; and media and screen time. Enrichment activities [$\eta_p^2 = 0.11$; $F_{(1,\ 44)} = 6.95$; $p < 0.02$] and social events [$\eta_p^2 = 0.10$; $F_{(1,\ 43)} = 7.26$; $p < 0.01$] significantly predicted self-directed EF, and play with others was marginally predictive [$\eta_p^2 = 0.05$; $F_{(1,\ 44)} = 3.42$; $p < 0.072$]. Interactions with age were not significant in these models, and were therefore excluded (p's > 0.2). No other classes of less-structured activities predicted verbal fluency performance.

We then considered whether the relationship between less-structured time and self-directed EF persisted when we excluded from our less-structured time composite measure, in sequential analyses:

(1) media and screen time (which might reflect passive, rather than self-directed leisure activity);
(2) activities within the less-structured time classification that may have included more structure than other such activities; and
(3) enrichment activities that may have yielded benefits specific to verbal fluency performance (rather than self-directed control, *per se*).

When media and screen time were excluded, less-structured time continued to demonstrate a positive relationship with self-directed EF [$\eta_p^2 = 0.06$; $F_{(1,\ 41)} = 5.23$; $p < 0.03$]. This finding persisted when we also excluded less-structured activities that may have included more structure than other such activities (e.g., board games played with a group; rule-based physical games such as golf and bowling; movies and performances; reading with others[6]) [$\eta_p^2 = 0.06$; $F_{(1,\ 43)} = 6.17$; $p < 0.02$]. As a final step, we also excluded visits to museums, aquariums, and zoos, which may have benefitted organization of semantic clusters on

[4]Aggregate within-measure scores were z-scored, then summed to create cross-measure composites.

[5]Social and enrichment events included only prior-week reporting, as these were not adequately identified in the annual less-structured time measure, which included only general activities (e.g., playing outdoor with friends) that could occur in many contexts.

[6]Here and in the following analysis, we also excluded all reading from our typical-activities measure, because this measure did not discriminate between reading alone and reading with others.

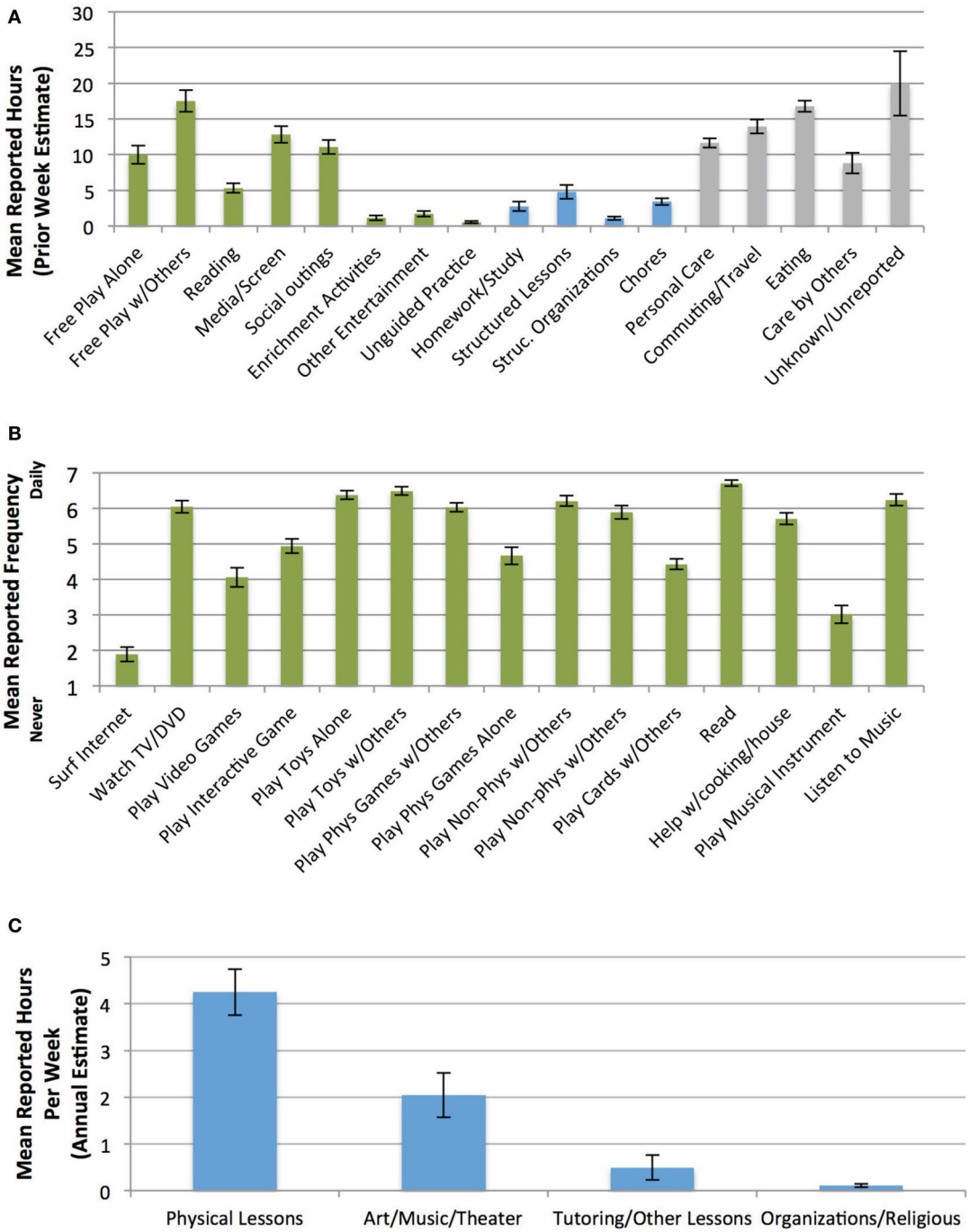


FIGURE 1 | Parent-reported child time use. (A) Activities in week prior to laboratory visit (green, less-structured; blue, structured; gray, other). **(B)** Typical less-structured activities (1, Never; 2, Less than once a month; 3, Once a month; 4, 2–3 times a month; 5, Once a week; 6, 2–3 times a week; 7, Daily). **(C)** Typical structured activities during a typical week (averaged across 4 seasons). Prior-week and typical measures of parent-reported child time use were correlated and combined into z-scored composite estimates of structured and less-structured time. For all figures, error bars indicate standard error of the mean.

the verbal fluency task (e.g., exposure to zoo animals may have helped to organize animal clusters, and thus yielded performance benefits). Using this fully-restricted measure of less-structured time, child time in less-structured activities continued to predict better self-directed EF [$\eta_p^2 = 0.06$; $F_{(1,\ 43)} = 6.23$; $p < 0.02$]. Interactions with age were significant and were included in each of these restricted analyses (all p's < 0.05).

We also explored whether participation in types of less-structured activities changed with age, and whether such changing patterns of time use could speak to the diminished link

between less-structured time and self-directed control in the oldest quartile of children in our sample. Media and screen time use was more prevalent in older children [$\eta_p^2 = 0.05$; $F_{(1, 61)} = 5.15$; $p < 0.03$]. Time spent in other categories of less-structured activities did not vary with age (p's > 0.2).

Structured time

Additional time in structured activities predicted marginally worse self-directed control [$\eta_p^2 = 0.06$; $F_{(1, 43)} = 3.57$; $p < 0.07$; **Figure 2B**; **Table 3**]. Again, self-directed EF was predicted by age [$\eta_p^2 = 0.13$; $F_{(1, 43)} = 4.43$; $p < 0.01$], and vocabulary [$\eta_p^2 = 0.08$; $F_{(1, 43)} = 5.02$; $p < 0.04$], and marginally predicted by household income [$\eta_p^2 = 0.05$; $F_{(1, 43)} = 3.50$; $p < 0.07$][7].

Exploratory analyses. We next examined whether the relationship between structured time and self-directed EF persisted when we excluded religious services and household chores, where children may have been supervised less often by adults, relative to other structured activities. Time in structured activities continued to predict worse self-directed EF when religious services and chores were excluded from the composite structured time measure [$\eta_p^2 = 0.06$; $F_{(1, 43)} = 4.28$; $p < 0.05$].

CHILD TIME USE AND EXTERNALLY-DRIVEN EF

No measure of child time predicted any aspect of externally-driven EF (**Figures 3A–D**). Specifically, child time spent in less-structured activities did not relate to performance on either externally-driven EF measure (Flanker conflict score: $p > 0.2$; AX-CPT proactive control score: $p > 0.8$). Similarly, time in structured activities was unrelated to externally-driven EF (Flanker conflict score: $p > 0.6$; AX-CPT proactive control score: $p > 0.3$)[8]. Males demonstrated better Flanker conflict scores than females [$\eta_p^2 = 0.10$; $F_{(1, 46)} = 4.64$; $p < 0.04$]. No other variables predicted externally-directed EF[9].

DISCUSSION

Our findings offer support for a relationship between the time children spend in less-structured and structured activities and the

[7]This finding was not driven by a negative correlation between composite time in structured activities and time in less-structured activities. The less-structured and structured time composites were not significantly related ($p > 0.8$).

Table 2 | Descriptive statistics for executive function, vocabulary, and time use measures (*N*'s = 65–67).

Measure	Mean (*SD*)
SELF-DIRECTED EF	
Verbal fluency combined switch score	10.13 (4.1)
EXTERNALLY-DRIVEN EF	
AX-CPT proactive control score	0.094 (0.12)
Flanker conflict score	164.5 (168.7)
Vocabulary: EVT standardized score	112.9 (9.4)
PRIOR WEEK CHILD TIME USE	
Structured hours	6.03 (5.9)
Less-structured hours	32.2 (14.2)
Typical child less-structured activities (combined score)	78.5 (8.8)
Seasonal child structured activities (annual hours)	91.5 (89.0)

[8]Although it is not a targeted measure of conflict resolution, overall accuracy across all trials on the Flanker task has also been tested in prior intervention work with children (Rueda et al., 2005; Fisher et al., 2011; Röthlisberger et al., 2012), and is what improved in two prior intervention studies targeting EF in this age group (Fisher et al., 2011; Röthlisberger et al., 2012). Overall accuracy did improve with age in our sample [$\eta_p^2 = 0.16$; $F_{(1, 47)} = 4.67$; $p < 0.04$], but was not predicted by any other variables (p's > 0.15).

[9]In separate analyses, we investigated whether the completeness of parent reporting of child time influenced observed relationships between child time use and EFs. For example, if parents who left fewer cells blank in the time use survey had children with higher self-directed EF, this could have contributed to the observed correlation between less-structured time and self-directed EF, since parents who left fewer cells blank might report more time in less-structured activities. However, completeness of reported time use did not affect the results: it showed no relationship with any aspect of EF performance (verbal fluency, AX-CPT, and Flanker p's > 0.3), and controlling for it did not change whether or not any findings were significant.

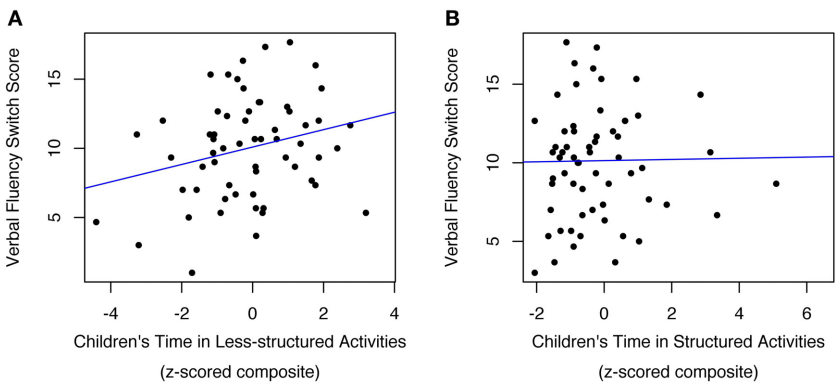


FIGURE 2 | Children's self-directed EF (as measured in Verbal Fluency) was predicted by more time spent in less-structured activities (A), and marginally predicted by less time spent in structured activities, although this relationship is not apparent because the figure does not capture how the effects of age, income, gender, and EVT were controlled for in all analyses (B). Outlying observations have been excluded [$N = 3$ in **(A)**; $N = 2$ in **(B)**].

Table 3 | Effects of age, gender, income, vocabulary and time use on child verbal fluency performance.

	Age, income, and gender			EVT, age, income, and gender			Less-structured time + age, income, gender, and EVT		
Variable	***β***	***t*** ^{b}	***P***	***β***	***t*** ^{b}	***P***	***β***	***t*** ^{b}	***P***
(Intercept)	9.99	22.48	<0.001***	9.95	22.13	<0.001***	9.94	22.67	<0.001***
Age (days)	0.008	1.79	<0.09	0.013	2.95	<0.01**	0.011	2.73	<0.01**
Gender (1 = female; −1 = male)	0.008	0.02	>0.9	−0.194	−0.42	>0.6	−0.392	−0.82	>0.4
Household income	0.743	3.17	<0.01**	0.301	1.13	>0.2	0.372	1.48	>0.1
Vocabulary (EVT)	–	–	–	0.177	2.78	<0.01**	0.142	2.51	<0.05*
Less-structured time	–	–	–	–	–	–	0.713	2.11	<0.05*
Less-structured time x age	–	–	–	–	–	–	–	–	–
Structured time	–	–	–	–	–	–	–	–	–
Model *F*-value		4.05			4.88			4.56	
Model Adjusted R^2		0.16			0.24			0.27	

	Less-structured time × age + income, gender, and EVT			Structured time + age, income, gender, and EVT		
Variable	***β***	***t*** ^{b}	***P***	***β***	***t*** ^{b}	***P***
(Intercept)	10.09	23.86	<0.001***	9.73	21.97	<0.001***
Age (days)	0.010	2.40	<0.05*	0.012	2.90	<0.01*
Gender (1 = female; −1 = male)	−0.375	−0.82	>0.4	−0.169	−0.36	>0.7
Household income	0.442	1.83	<0.08	0.487	1.87	<0.07
Vocabulary (EVT)	0.120	2.19	<0.05*	0.128	2.24	<0.05*
Less-structured time	0.854	2.61	<0.05*	–	–	–
Less-structured time x age	−0.008	−2.34	<0.05*	–	–	–
Structured time	–	–	–	−0.596	−1.89	<0.07
Model *F*-value		5.10			4.34	
Model Adjusted R^2		0.33			0.26	

Age, income, EVT scores, and less-structured and structured time composite scores are mean-centered. For each model, observations where Cook's D > 3 standard deviations above the mean were identified and removed. $N_{Model1} = 45$; $N_{Models2–3} = 44$; $N_{Models4–5} = 43$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

development of self-directed executive function. When considering our entire participant sample, children who spent more time in less-structured activities displayed better self-directed control, even after controlling for age, verbal ability, and household income. By contrast, children who spent more time in structured activities exhibited poorer self-directed EF, controlling for the same factors. The observed relationships between time use and EF ability were specific to self-directed EF, as neither structured nor less-structured time related to performance on externally-driven EF measures. These findings represent the first demonstration that time spent in a broad range of less-structured activities outside of formal schooling predicts goal-directed behaviors not explicitly specified by an adult, and that more time spent in structured activities predicts poorer such goal-directed behavior. Consistent with Vygotskian developmental theory and programs that build on that theory, such as Tools of the Mind, less-structured time may uniquely support the development of self-directed control by affording children with additional practice in carrying out goal-directed actions using internal cues and reminders. That is, less-structured activities may give children more self-directed opportunities. From this perspective, structured time could slow the development of self-directed control, since adults in such scenarios can provide external cues and reminders about what should happen, and when.

Surprisingly, the relationship between less-structured time and self-directed control changed with age in our participant sample, such that less-structured time predicted self-directed control in all but the oldest quartile of participants. This interaction between less-structured time and age was reliably observed across increasingly restrictive measures of less-structured time. One interpretation is that most but not all age groups within our sample spent their less-structured time in activities that encourage the development of self-directed control. Indeed, despite a relatively limited age range, our sample demonstrated differences in the content of less-structured time across 6–7 years of age, with older children spending more time engaged in media and screen activities. However, time spent in unguided practice, enrichment outings, and some forms of play was the main driver of the relationship between less-structured time and self-directed control in our data, and time spent in such activities did not change as a

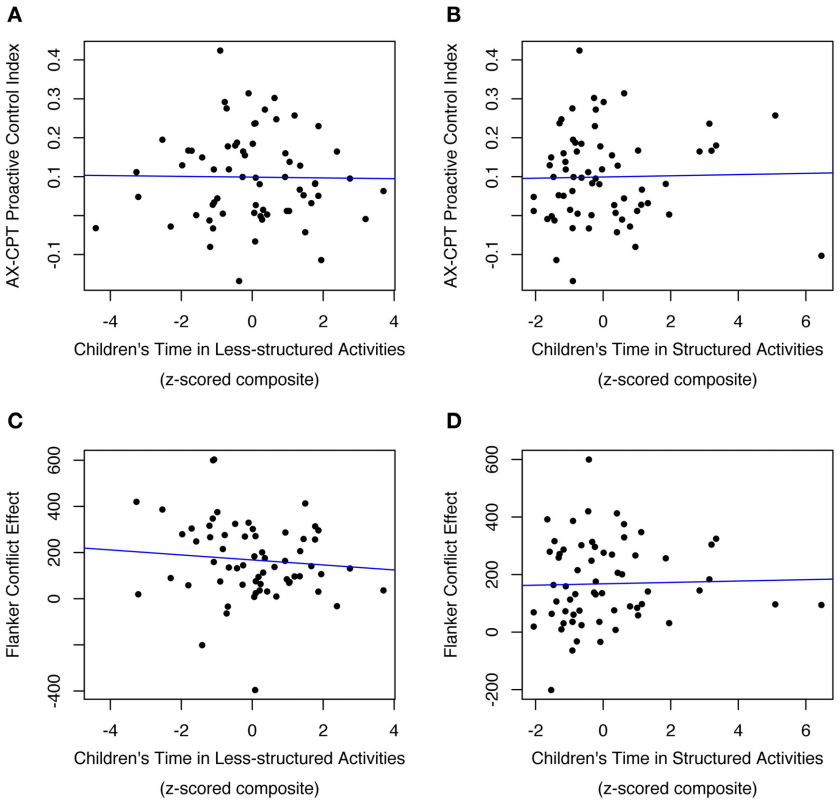


FIGURE 3 | Children's externally-driven EF (as measured in AX-CPT and Flanker) was not predicted by their time spent in either less-structured activities (A,C) or structured activities (B,D). Outlying observations have been excluded [$N = 1$ in **(A,B)**; $N = 2$ in **(C,D)**].

function of age. Another possibility is that children who have less developed self-directed control are more likely to benefit from less-structured time (in the same way that some interventions show the greatest benefits to children who show the worst initial performance, Connor et al., 2010; Diamond and Lee, 2011; cf. Bierman et al., 2008), such that the oldest and most advanced quartile of participants showed the least benefit.

While promising, it will be important for the present findings to be replicated and extended to address a number of limitations. For example, our sample came primarily from an affluent, suburban sample. This sample nonetheless included a broad enough range of incomes that income was predictive of self-directed EF, and the relationship between less-structured time and self-directed EF held even when controlling for income. However, less-structured time may be especially beneficial to children in safe, quiet, resource-rich environments, so it will be important to test whether it differentially relates to self-direction in more impoverished environments. In addition, although the current test of the relationship between less-structured time and self-directed EFs emerged from a targeted hypothesis, we conducted multiple post-hoc exploratory analyses to explore the relationship between specific activities and self-directed control, which are not ideal conditions for statistical inference.

Another limitation of the present study relates to our constructions of less-structured and structured time, which are imprecise, and most likely fail to capture important differences across activities. The broad, standardized definitions of structured and less-structured time adopted in this study (e.g., Meeks and Mauldin, 1990) ignore differences in the degree of independence that children experience within and across activities. In the present study, trips to museums, libraries, and sporting events are each classified as less-structured, but may vary in relative structure. That is, a typical library visit, where children may select their own sections to browse and books to check out, may involve much less structure (and more self-directed time) than a typical sporting event, where attention is largely directed toward the action on the field or court. Similarly, although any activity within the category of "media and screen time" counts as less-structured time, this category includes activities that range from passive movie-watching to self-directed internet searches to more structured video games. Even those activities that seem less-structured by definition, such as free play, can quickly become more structured when adults, older siblings, or peers impose additional rules or criteria. Indeed, many programmatic interventions have highlighted the importance of some structure to improve the quality of children's play and other learning experiences, and produce benefits (Schweinhart et al., 2005; Lillard and Else-quest, 2006; Diamond et al., 2007; Heckman et al., 2010; Lillard, 2012).

We note however, that even though our classification system based on the existing literature does not capture these variations

in exactly how structured various activities are, our primary finding of the relationship between less-structured time and self-directed EF holds across analyses dropping potentially more difficult-to-interpret classifications (e.g., media and screen time, various games, movies and performances, and visits to museums, aquariums and zoos). To generate a more precise estimate of the amount of time children spend pursuing activities in a self-directed way, one would ideally assess child time directly, possibly by supplementing parent-reported child time use data with direct observation. One possibility along these lines could be to employ experience sampling techniques (Miller, 2012), where parents are frequently queried (via cell phone or another mobile device) throughout the day and asked to provide specific detail about their child's activities in the moment. Such methods would also minimize the need to rely on a parent's memory for their child's daily activities and experiences. We view our work as providing an important starting point for this kind of more time-intensive study of children's time outside of formal schooling and its relationship to their self-directed EF.

In addition, although we have identified links between child time use and self-directed EF, we are unable to draw firm conclusions about whether the observed relationships were driven by activities occurring in the week preceding the test session (as has been observed in other domains, e.g., Berns et al., 2013; Mackey et al., 2013), activities occurring over a longer period, or some combination. We used composite measures incorporating both recent and more distal/typical experiences, given that these measures were correlated and in an attempt to maximize the accuracy and reliability of parental estimates. We can test which one is more predictive of self-directed EF, recent or more distal/typical experiences, but it is difficult to make strong claims based on such analyses. For example, when examining less-structured activities and self-directed EF, we find that recent experiences predict self-directed EF [$F_{(1,\ 60)} = 6.10$; $p < 0.02$], but typical experiences do not ($p > 0.6$). This finding could reflect the greater importance of recent experiences, or it could reflect the greater precision of the time-diary measure, which indexes recent experiences but is also representative of more distal/typical experiences[10]. Similarly, when examining structured activities and self-directed EF, we find that neither recent nor annual experiences alone predict self-directed EF (p's > 0.2). This finding could reflect the importance of the combination of recent and distal experiences, or simply the greater robustness of using a composite measure. Therefore, while we have posited that less-structured experiences allow children to practice self-directed, goal-oriented behavior, producing benefits over time, we cannot discount the possibility that observed linkages may have been driven by recent experiences which increased self-directed behavior. In either scenario, regular participation in less-structured activities would yield benefits.

Future investigations of the relationship between self-directed control and less-structured time would also benefit from the inclusion of additional measures of self-directed control, which more closely approximate real-world child behaviors. This process may benefit from the development and validation of new measures of self-directed control in children. Establishing effects using tasks tapping other forms of self-direction would also ensure generalizability. For instance, in the present study, time in less-structured activities such as family outings may have benefitted verbal fluency performance in a specific way, by fostering the development of more well-organized semantic networks, rather than by more generally improving children's abilities to generate their own rules for how and when to employ EFs to achieve their goals. This alternative account cannot explain the full pattern of results in the link between less-structured time and self-directed EF (e.g., the fact that this link persists when enrichment activities are excluded, and other less-structured categories such as unguided practice and play predict self-directed EF); however, a broader range of measures could provide a more robust and generalizable assessment of self-directed EF.

The findings of the current study are consistent with previous research in showing a link between children's experiences and EF (Lillard and Else-quest, 2006; Diamond et al., 2007; Bierman et al., 2008; Holmes et al., 2009; Bergman Nutley et al., 2011; Diamond, 2012; Röthlisberger et al., 2012; Zelazo and Lyons, 2012; Titz and Karbach, 2014). However, while the current study found specific effects of time use on self-directed but not externally-driven EF, previous research found effects of training and preschool interventions on externally-driven EF (e.g., see discussion in Diamond, 2012), but did not evaluate self-directed EF. There are several possible reasons for this discrepancy. First, previous training studies that have shown benefits for externally-driven EF have specifically trained children on aspects of externally-driven EF (e.g., working memory span tasks; e.g., Holmes et al., 2009; Bergman Nutley et al., 2011). Likewise, while preschool and other interventions include a wide variety of experiences, they likely include considerable practice with externally-driven EF. In contrast, we hypothesize that less-structured time primarily affords children practice with self-directed EF, and thus may not transfer to improving externally-driven EF. Second, it is possible that differences between the current versus previous studies could be accounted for by differences between the externally-driven EF tasks they employed. Many previous studies that have found effects of interventions on externally-driven EF used task-switching or working memory span tasks (e.g., Lillard and Else-quest, 2006; Diamond et al., 2007; Bierman et al., 2008; Holmes et al., 2009, 2010; Thorell et al., 2009; Bergman Nutley et al., 2011; Röthlisberger et al., 2012), whereas the current study used tasks assessing proactive control (AX-CPT) and conflict resolution (Flanker). It may be that specific aspects of externally-driven EF are more sensitive to children's experiences, or that

[10] Recent less-structured experiences also predict self-directed EF when controlling for parent-reported typicality of the prior week [$\eta_p^2 = 0.02$; $F_{(1,\ 57)} = 9.12$; $p < 0.004$; $M_{typ} = 4$; $SD = 2.05$; range = 1–7], and there is no interaction between less-structured experiences and typicality in predicting self-directed EF ($p > 0.8$). These findings might suggest that the prior week's experience is predictive separate from the extent to which it reflects typical/distal experiences. However, this interpretation rests on the validity and sensitivity of the typicality measure, which is unknown. Parent-reported typicality is at least internally consistent with parent-reported time use. Specifically, recent less-structured experiences predicted typical/distal experiences when parent-reported typicality of the prior week was high [$M_{typ} = 6$; $\eta_p^2 = 0.03$; $F_{(1,\ 60)} = 5.81$; $p < 0.02$], but not when typicality of the prior week was low ($M_{typ} = 2$; $p > 0.9$), yielding a marginally significant interaction [$M_{typ} = 2$; $\eta_p^2 = 0.04$; $F_{(1,\ 60)} = 2.92$; $p < 0.093$].

specific tasks are more sensitive to individual differences in general due to better psychometric properties[11]. Future research using a more comprehensive battery of EF tasks could address these possibilities.

Another key difference between our study and such prior research is the correlational nature of our study, which supports at least two alternatives to the interpretation that how children spend their leisure time shapes their EF. First, children with better self-directed EFs may engage in (or be encouraged to engage in) less-structured activities more often. Likewise, children with poorer self-directed control may be more likely to engage in structured activities. Alternatively, the observed relationship between less-structured time and self-directed control may be driven by a third, unmeasured variable. Although we have attempted to control for some characteristics that might influence both time spent in less-structured activities and verbal fluency, such as household income, we have not controlled for other possibilities, such as parent EF and child's fluid intelligence (which we did not assess). However, we did control for child vocabulary (an index of crystallized intelligence), which may serve as a proxy for fluid intelligence in testing relationships with EF, given that EF fully mediates the correlation between crystallized and fluid intelligence in 7-years-old (Brydges et al., 2012)[12]. Moreover, such factors might be expected to predict both children's self-directed EF and their externally-driven EF (Ardila et al., 2000; Mahone et al., 2002; Kalkut et al., 2009), and so seem unlikely to explain why less-structured time predicts only the former. Similar issues have been raised in interpreting links observed between children's EF and pretend play: rather than reflecting a uniquely causal role for pretend play in EF, EF may instead play a causal role in supporting pretend play, or pretend play may be one of many activities promoting EF development in young children (Lillard et al., 2013).

An important direction for future work lies in establishing the directionality of relationships between child time use and self-directed EF, through experimental manipulation. Longitudinal studies could provide the first step toward establishing directionality. Specifically, if time spent in less-structured activities prospectively predicts change in self-directed EF, this would suggest that less-structured time may play a causal role in the development of self-directed EF. If, on the other hand, self-directed EF prospectively predicts changes in the amount of time children spend in less-structured activities, this would suggest that self-directed EF may play a causal role in children's time use (e.g., because parents might allow children with strong self-directed EF skills to play with less supervision). While such longitudinal studies could thus provide important information about temporal precedence, this information is not sufficient evidence of causality (e.g., additional unmeasured variables could actually be the causal factors). Thus, future research using experimental manipulations of time spent in less-structured activities is necessary to definitively test causality. One approach would be to attempt to randomly assign children to more structured or less structured environments, such as summer camps, where child activities could be carefully monitored via regular sampling of staff and/or on-site observation. Although this kind of work is ambitious, and poses challenges, it could be used to inform more targeted laboratory-based training studies.

Finally, we hope that future explorations of the relationship between child time use and developing self-directed EFs will inform a wider question: specifically, whether societal shifts in child time use over the past 50 years have influenced development. Hours formerly devoted to less-structured, social play have been replaced by media time (Vandewater et al., 2007; Bavelier et al., 2010; Hofferth, 2010; Johnson, 2010), and structured, adult-led activities (Hofferth and Sandberg, 2001a; Larson, 2001; Bianchi et al., 2006). Some commentators have warned that these changes have been to the detriment of children (e.g., Ginsburg, 2007; Milteer and Ginsburg, 2012). Others have argued that children benefit more from regular skill practice in structured settings (e.g., Chua, 2011; Ramdass and Zimmerman, 2011). Our findings indicate that during children's time outside of formal schooling, participation in less structured activities may benefit the development of self-directed EFs, while participation in structured activities may hinder the development of self-directed EFs. Thorough testing of this hypothesis remains an important direction for future work.

AUTHOR CONTRIBUTIONS

Jane E. Barker, Andrei D. Semenov, and Yuko Munakata contributed to the development of the study hypothesis. All authors contributed to study design. Jane E. Barker performed the data analysis and drafted the manuscript with input from Yuko Munakata. Critical revisions were contributed by Hannah R. Snyder, Laura Michaelson, and Yuko Munakata. All authors discussed the results, implications, and literature, and approved the final version of the manuscript for submission.

ACKNOWLEDGMENTS

The authors wish to thank Julia Stadele for her assistance in coding verbal fluency data and coordinating research subjects, Joe Brill for his assistance in coding time diary data, and Ryan Guild for his helpful comments on manuscript revisions. This research was supported by a grant from the National Institute of Child Health and Human Development (RO1 HD37163). Publication of this article was funded by the University of Colorado Boulder Libraries Open Access Fund.

[11] For example, some EF-interventions have not improved performance on the Flanker task in this age group (Rueda et al., 2005, 2012; see also Diamond et al., 2007, which introduced switching demands that did show effects of intervention, and included only incongruent trials so that a standard conflict score could not be computed). The Flanker task can be sensitive to minor variations in stimulus parameters (Paquet, 2001) and intervention dosage in adults (Liu-Ambrose et al., 2012). Failures to find effects of interventions have also been attributed in part to the task's sensitivity to practice effects in pre-post measure designs (as discussed in Rueda et al., 2012), which are not an issue in the present study.

[12] We also note that there is ongoing debate regarding the inappropriateness of IQ as a control in models of cognitive processes (Dennis et al., 2009).

REFERENCES

Abwender, D. A., Swan, J. G., Bowerman, J. T., and Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: review and comparison of several scoring methods. *Assessment* 8, 323–338. doi: 10.1177/107319110100800308

Agnew, R., and Petersen, D. M. (1989). Leisure and delinquency. *Soc. Probl.* 36, 332–350. doi: 10.2307/800819

Albertson, K., and Shore, C. (2009). Holding in mind conflicting information: pretending, working memory, and executive control. *J. Cogn. Dev.* 9, 390–410. doi: 10.1080/15248370802678240

Ardila, A., Pineda, D., and Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Arch. Clin. Neuropsychol.* 15, 31–36. doi: 10.1093/arclin/15.1.31

Bavelier, D., Green, C. S., and Dye, M. W. G. (2010). Children, wired: for better and worse. *Neuron* 67, 692–701. doi: 10.1016/j.neuron.2010.08.035

Bergman Nutley, S., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., and Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: a controlled, randomized study. *Dev. Sci.* 14, 591–601. doi: 10.1111/j.1467-7687.2010.01022.x

Berns, G. S., Blaine, K., Prietula, M. J., and Pye, B. E. (2013). Short- and long-term effects of a novel on connectivity in the brain. *Brain Connect.* 3, 590–600. doi: 10.1089/brain.2013.0166

Best, J. R., Miller, P. H., and Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learn. Individ. Differ.* 21, 327–336. doi: 10.1016/j.lindif.2011.01.007

Bianchi, S. M., Robinson, J. P., and Milkie, M. A. (2006). *Changing Rhythms of American Family Life*. ASA Rose Series. New York, NY: Sussell Sage.

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., and Domitrovich, C. E. (2008). Executive functions and school readiness intervention: impact, moderation, and mediation in the Head Start REDI program. *Dev. Psychopathol.* 20, 821–843. doi: 10.1017/S0954579408000394

Birn, R. M., Kenworthy, L., Case, L., Caravella, R., Jones, T. B., Bandettini, P. A., et al. (2010). Neural systems supporting lexical search guided by letter and semantic category cues: a self-paced overt response fMRI study of verbal fluency. *Neuroimage* 49, 1099–1107. doi: 10.1016/j.neuroimage.2009.07.036

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Blair, C., Granger, D., and Peters Razza, R. (2005). Cortisol reactivity is positively related to executive function in preschool children attending Head Start. *Child Dev.* 76, 554–567. doi: 10.1111/j.1467-8624.2005.00863.x

Bodrova, E. (2003). Vygotsky and Montessori: one dream, two visions. *Montessori Life* 15, 30–33.

Bodrova, E., and Leong, D. J. (2007). *Tools of the Mind*. Columbus, OH: Pearson.

Braver, T. S., Gray, J. R., and Burgess, G. C. (2007). "Explaining the many varieties of working memory variation: dual mechanisms of cognitive control," in *Variation in Working Memory,* eds A. Conway, C. Jarrold, M. Kane, A. Miyake, and J. Towse (New York, NY: Oxford University Press), 76–106.

Brocki, K. C., and Bohlin, G. (2004). Executive functions in children aged 6 to 13: a dimensional and developmental study. *Dev. Neuropsychol.* 26, 571–593. doi: 10.1207/s15326942dn2602_3

Bryck, R. L., and Mayr, U. (2005). On the role of verbalization during task set selection: switching or serial order control? *Mem. Cognit.* 33, 611–623. doi: 10.3758/BF03195328

Brydges, C. R., Reid, C. L., Fox, A. M., and Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence* 40, 458–469. doi: 10.1016/j.intell.2012.05.006

Cameron, C. E., Brock, L. L., Murrah, W. M., Bell, L. H., Worzalla, S. L., Grissmer, D., et al. (2012). Fine motor skills and executive function both contribute to kindergarten achievement. *Child Dev.* 83, 1229–1244. doi: 10.1111/j.1467-8624.2012.01768.x

Carlson, S. M., White, R. E., and Davis-Unger, A. (2014). Evidence for a relation between executive function and pretense representation in preschool children. *Cogn. Dev.* 29, 1–16. doi: 10.1016/j.cogdev.2013.09.001

Chatham, C. H., Frank, M. J., and Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5529–5533. doi: 10.1073/pnas.0810002106

Chevalier, N., Wiebe, S. A., Huber, K. L., and Espy, K. A. (2011). Switch detection in preschoolers' cognitive flexibility. *J. Exp. Child Psychol.* 109, 353–370. doi: 10.1016/j.jecp.2011.01.006

Chrysikou, E. G., and Thompson-Schill, S. L. (2011). Dissociable brain states linked to common and creative object use. *Hum. Brain Mapp.* 32, 665–675. doi: 10.1002/hbm.21056

Chua, A. (2011). *Battle Hymn of the Tiger Mother*. New York, NY: Penguin.

Connor, C. M., Ponitz, C. C., Phillips, B. M., Travis, Q., Glasney, S., and Morrison, F. J. (2010). First graders' literacy and self-regulation gains: the effect of individualizing student instruction. *J. Sch. Psychol.* 48, 433–455. doi: 10.1016/j.jsp.2010.06.003

Costafreda, S. G., Fu, C. H., Lee, L., Everitt, B., Brammer, M. J., and David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. *Hum. Brain Mapp.* 27, 799–810. doi: 10.1002/hbm.20221

Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., and Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *J. Int. Neuropsychol. Soc.* 15, 331–343. doi: 10.1017/S1355617709090481

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Curr. Dir. Psychol. Sci.* 21, 335–341. doi: 10.1177/0963721412453722

Diamond, A., Barnett, W. S., Thomas, J., and Munro, S. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4-12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529

D'Ostilio, K., and Garraux, G. (2012). Dissociation between unconscious motor response facilitation and conflict in medial frontal areas. *Eur. J. Neurosci.* 35, 332–340. doi: 10.1111/j.1460-9568.2011.07941.x

Drane, D. L., Lee, G. P., Cech, H., Huthwaite, J. S., Ojemann, G. A., Ojemann, J. G., et al. (2006). Structured cueing on a semantic fluency task differentiates patients with temporal versus frontal lobe seizure onset. *Epilepsy Behav.* 9, 339–344. doi: 10.1016/j.yebeh.2006.06.010

Eccles, J. S., and Barber, B. (1999). Student council, volunteering, basketball, or marching band: what kind of extracurricular involvement matters? *J. Adolesc. Res.* 14, 10–34. doi: 10.1177/0743558499141003

Elias, C. L., and Berk, L. E. (2002). Self-regulation in young children: is there a role for sociodramatic play? *Early Child. Res. Q.* 17, 216–238. doi: 10.1016/S0885-2006(02)00146-1

Ellamil, M., Dobson, C., Beeman, M., and Christoff, K. (2012). Evaluative and generative modes of thought during the creative process. *Neuroimage* 59, 1783–1794. doi: 10.1016/j.neuroimage.2011.08.008

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128, 309–331. doi: 10.1037/0096-3445.128.3.309

Eriksen, C. W., and Schultz, D. W. (1979). Information processing in visual search: a continuous flow conception and experimental results. *Percept. Psychophys.* 25, 249–263. doi: 10.3758/BF03198804

Farah, M. J., Shera, D. M., Savage, J. H., Betancourt, L., Giannetta, J. M., Brodsky, N. L., et al. (2006). Childhood poverty: specific associations with neurocognitive development. *Brain Res.* 1110, 166–174. doi: 10.1016/j.brainres.2006.06.072

Fisher, A., Boyle, J. M., Paton, J. Y., Tomporowski, P., Watson, C., McColl, J. H., et al. (2011). Effects of a physical education intervention on cognitive function in young children: randomized controlled pilot study. *BMC Pediatr.* 11:97. doi: 10.1186/1471-2431-11-97

Fletcher, A. C., Nickerson, P., and Wright, K. L. (2003). Structured leisure activities in middle childhood: links to well-being. *J. Community Psychol.* 31, 641–659. doi: 10.1002/jcop.10075

Forstmann, B. U., Brass M., Koch, I., and von Cramon, D. Y. (2005). Internally generated and directly cued task sets: an investigation with fMRI. *Neuropsychologia* 43, 943–952. doi: 10.1016/j.neuropsychologia.2004.08.008

Gathercole, S. E., Pickering, S. J., Ambridge, B., and Wearing H. (2004). The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* 40, 177–190. doi: 10.1037/0012-1649.40.2.177

Ginsburg, K. R. (2007). The importance of play in promoting healthy child development and maintaining strong parent-child bonds. *Pediatrics* 119, 182–191. doi: 10.1542/peds.2006-2697

Groborz, M., and Nęcka, E. (2003). Creativity and cognitive control: explorations of generation and evaluation skills. *Creat. Res. J.* 15, 183–197. doi: 10.1080/10400419.2003.9651411

Heckman, J. J., Moon. S. H., Pinto, R. P., Savelyev, A., and Yavitz, A. Q., (2010). The rate of return to the highscope perry preschool program. *J. Public Econ.* 94, 114–128. doi: 10.1016/j.jpubeco.2009.11.001

Henry, J. D., and Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology* 18, 284. doi: 10.1037/0894-4105.18.2.284

Hirshorn, E. A., and Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: neural correlates of switching during verbal fluency. *Neuropsychologia* 44, 2547–2557. doi: 10.1016/j.neuropsychologia.2006.03.035

Hofferth, S. L. (2010). Home media and children's achievement and behavior. *Child Dev.* 81, 1598–1619. doi: 10.1111/j.1467-8624.2010.01494.x

Hofferth, S. L., and Sandberg, J. F. (2001a). Changes in American children's time, 1981-1997. *Adv. Life Course Res.* 6, 193–229. doi: 10.1016/S1040-2608(01)80011-3

Hofferth, S. L., and Sandberg, J. F. (2001b). How American children spend their time. *J. Marriage Fam.* 63, 295–308. doi: 10.1111/j.1741-3737.2001.00295.x

Holmes, J., Gathercole, S. E., and Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Dev. Sci.* 12, 9–15. doi: 10.1111/j.1467-7687.2009.00848.x

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., and Elliott, J. G. (2010). Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. *Appl. Cogn. Psychol.* 24, 827–836. doi: 10.1002/acp.1589

Hughes, C., Ensor, R., Wilson, A., and Graham, A. (2009). Tracking executive function across the transition to school: a latent variable approach. *Dev. Neuropsychol.* 35, 20–36. doi: 10.1080/87565640903325691

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Iudicello, J. E., Kellogg, E. J., Weber, E., Smith, C., Grant, I., Drane, D. L., et al. (2012). Semantic cueing improves category verbal fluency in persons living with HIV infection. *J. Neuropsychiatr. Clin. Neurosci.* 24, 183–190. doi: 10.1176/appi.neuropsych.11100301

Jacques, S., and Zelazo, P. D. (2001). Developmental Neuropsychology The Flexible Item Selection Task (FIST): a measure of executive function in preschoolers. *Dev. Neuropsychol.* 20, 573–591. doi: 10.1207/S15326942DN2003_2

Jarosz, A. F., Colflesh, G. J., and Wiley, J. (2012). Uncorking the muse: alcohol intoxication facilitates creative problem solving. *Conscious. Cogn.* 21, 487–493. doi: 10.1016/j.concog.2012.01.002

Johnson, G. M. (2010). Children's Internet use at home and school: patterns and profiles. *J. Early Child. Res.* 8, 282–293. doi: 10.1177/1476718X10379783

Kalkut, E. L., Han, S. D., Lansing, A. E., Holdnack, J. A. and Delis, D. C. (2009). Development of set-shifting ability from late childhood through early adulthood. *Arch. Clin. Neuropsychol.* 24, 565–574. doi: 10.1093/arclin/acp048

Kave, G., Kigel, S., and Kochva, R. (2008). Switching and clustering in verbal fluency tasks throughout childhood. *J. Clin. Exp. Neuropsychol.* 30, 349–359. doi: 10.1080/13803390701416197

Kelly, R., Hammond, S., Dissanayake, C., and Ihsen, E. (2011). The relationship between symbolic play and executive function in young children. *Australas. J. Early Child.* 36, 1–27.

Koren, R., Kofman, O., and Berger, A. (2005). Analysis of word clustering in verbal fluency of school-aged children. *Arch. Clin. Neuropsychol.* 20, 1087–1104. doi: 10.1016/j.acn.2005.06.012

Larson, R. W. (2001). How US children and adolescents spend time: what it does (and doesn't) tell us about their development. *Curr. Dir. Psychol. Sci.* 10, 160–164. doi: 10.1111/1467-8721.00139

Larson, R. W., and Verma, S. (1999). How children and adolescents spend time across the world: work, play, and developmental opportunities. *Psychol. Bull.* 125, 701–736. doi: 10.1037/0033-2909.125.6.701

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *J. Exp. Psychol.* 21, 451–468. doi: 10.1037/0096-1523.21.3.451

Lie, C., Specht, K., Marshall, J., and Fink, G. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *Neuroimage* 30, 1038–1049. doi: 10.1016/j.neuroimage.2005.10.031

Lillard, A., and Else-quest, N. (2006). Evaluating montessori education. *Science* 313, 1893–1894. doi: 10.1126/science.1132362

Lillard, A. S. (2012). Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. *J. Sch. Psychol.* 50, 379–401. doi: 10.1016/j.jsp.2012.01.001

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., and Palmquist, C. M. (2013). The impact of pretend play on children's development: a review of the evidence. *Psychol. Bull.* 139, 1–34. doi: 10.1037/a0029321

Liu-Ambrose, T., Nagamatsu, L. S., Voss, M. W., Khan, K. M., and Handy, T. C. (2012). Resistance training and functional plasticity of the aging brain: a 12-month randomized controlled trial. *Neurobiol. Aging* 33, 1690–1698. doi: 10.1016/j.neurobiolaging.2011.05.010

Mackey, A. P., Singley A. T. M., and Bunge S. A. (2013). Intensive reasoning training alters patterns of brain connectivity at rest. *J. Neurosci.* 33, 4796–4803. doi: 10.1523/JNEUROSCI.4141-12.2013

Mahone, M. E., Hagelthorn, K. M., Cutting, L. E., Schuerholz, L. J., Pelletier, S. F., Rawlins, C., et al. (2002). Effects of IQ on executive function measures in children with ADHD. *Child Neuropsychol.* 8, 52–65. doi: 10.1076/chin.8.1.52.8719

Mahoney, J. L., and Stattin, H. (2000). Leisure activities and adolescent antisocial behavior: the role of structure and social context. *J. Adolesc.* 23, 113–127. doi: 10.1006/jado.2000.0302

McAuley, T., Christ, S. E., and White, D. A. (2011). Mapping the development of response inhibition in young children using a modified day-night task. *Dev. Neuropsychol.* 36, 539–551. doi: 10.1080/87565641.2010.549871

McDermott, J. M., Pérez-Edgar, K., and Fox, N. A. (2007). Variations of the flanker paradigm: assessing selective attention in young children. *Behav. Res. Methods* 39, 62–70. doi: 10.3758/BF03192844

Meeks, C. B., and Mauldin, T. (1990). Children's time in structured and unstructured leisure activities. *Lifestyles Fam. Econ. Issues* 11, 257–281. doi: 10.1007/BF00987003

Miller, G. (2012). The smartphone psychology manifesto. *Perspect. Psychol. Sci.* 7, 221–237. doi: 10.1177/1745691612441215

Miller, M. R., Müller, U., Giesbrecht, G. F., Carpendale, J. I. M., and Kerns, K. A. (2013). The contribution of executive function and social understanding to preschoolers' letter and math skills. *Cogn. Dev.* 28, 331–349. doi: 10.1016/j.cogdev.2012.10.005

Milteer, R. M., and Ginsburg, K. R. (2012). The importance of play in promoting healthy child development and maintaining strong parent-child bond: focus on children in poverty. *Pediatrics* 129, 204–213. doi: 10.1542/peds.2011-2953

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108

Montessori, M. M. Jr. (1976). *Education for Human Development: Understanding Montessori.* New York, NY: Schocken.

Munakata, Y., Snyder, H. R., and Chatham, C. H. (2012). Developing cognitive control: three key transitions. *Curr. Dir. Psychol. Sci.* 21, 71–77. doi: 10.1177/0963721412436807

Noble, K. G., McCandliss, B. D., and Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Dev. Sci.* 10, 464–480. doi: 10.1111/j.1467-7687.2007.00600.x

Noble, K. G., Norman, M. F., and Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Dev. Sci.* 8, 74–87. doi: 10.1111/j.1467-7687.2005.00394.x

Osgood, D. W., Anderson, A. L., and Shaffer, J. N. (2005). "Unstructured leisure in the after-school hours," in *Organized Activities as Contexts of Development: Extracurricular Activities, After-school and Community Programs,* eds J. L. Mahoney, R. W. Larson, J. S. Eccles (Mahwah, NJ: Erlbaum), 45–64.

Paquet, L. (2001). Eliminating flanker effects and negative priming in the flankers task: evidence for early selection. *Psychon. Bull. Rev.* 8, 301–306. doi: 10.3758/BF03196165

Raboutet, C., Sauzéon, H., Corsini, M. M., Rodrigues, J., Langevin, S., and N'Kaoua, B. (2010). Performance on a semantic verbal fluency task across time: dissociation between clustering, switching, and categorical exploitation processes. *J. Clin. Exp. Neuropsychol.* 32, 268–280. doi: 10.1080/13803390902984464

Ramdass, D., and Zimmerman, B. J. (2011). Developing self-regulation skills: the important role of homework. *J. Adv. Acad.* 22, 194–218. doi: 10.1177/1932202X1102200202

Randolph, C., Braun, A. R., Goldberg, T. E., and Chase, T. N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: dissociation of

storage and retrieval failures. *Neuropsychology* 7, 82–88. doi: 10.1037/0894-4105.7.1.82

Raver, C. C., Blair, C., and Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: new perspectives on models of differential susceptibility. *Dev. Psychol.* 49, 292–304. doi: 10.1037/a0028343

Ridderinkhof, K. R., and van der Molen, M. W. (1995). A psychophysiological analysis of developmental differences in the ability to resist interference. *Child Dev.* 66, 1040–1056. doi: 10.2307/1131797

Röthlisberger, M., Neuenschwander, R., Cimeli, P., Michel, E., and Roebers, C. M. (2012). Improving executive function in 5- and 6-year-olds: evaluation of a small group intervention in prekindergarten and kindergarten children. *Infant Child Dev.* 21, 411–429. doi: 10.1002/icd.752

Rueda, M. R., Checa, P., and Cómbita, L. M. (2012). Enhanced efficiency of the executive attention network after training in preschool children: immediate changes and effects after two months. *Dev. Cogn. Neurosci.* 2, S192–S204. doi: 10.1016/j.dcn.2011.09.004

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., and Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14931–14936. doi: 10.1073/pnas.0506897102

Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., and Claverie, B. (2004). Verbal fluency output in children aged 7–16 as a function of the production criterion: qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain Lang.* 89, 192–202. doi: 10.1016/S0093-934X(03)00367-5

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., and Nores, M. (2005). *Lifetime Effects: The High/Scope Perry Preschool study Through age 40* (Monographs of the High/Scope Educational Research Foundation, 14). Ypsilanti, MI: High/Scope Press.

Shipstead, Z., Redick, T. S., and Engle, R. W. (2012). Is working memory training effective? *Psychol. Bull.* 138, 628–654. doi: 10.1037/a0027473

Smidts, D. P., Jacobs, R., and Anderson, V. (2004). The Object Classification Task for Children (OCTC): a measure of concept generation and mental flexibility in early childhood. *Dev. Neuropsychol.* 26, 385–401. doi: 10.1207/s15326942dn2601_2

Snyder, H. R., and Munakata, Y. (2010). Becoming self-directed: abstract representations support endogenous flexibility in children. *Cognition* 116, 155–167. doi: 10.1016/j.cognition.2010.04.007

Snyder, H. R., and Munakata, Y. (2013). So many options, so little control: abstract representations can reduce selection demands to increase children's self-directed flexibility. *J. Exp. Child Psychol.* 116, 659–673. doi: 10.1016/j.jecp.2013.07.010

St Clair-Thompson, H. L., and Gathercole, S. E. (2006). Executive functions and achievements in school: shifting, updating, inhibition, and working memory. *Q. J. Exp. Psychol.* 59, 745–759. doi: 10.1080/17470210500162854

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., and Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12, 106–113. doi: 10.1111/j.1467-7687.2008.00745.x

Titz, C., and Karbach, J. (2014). Working memory and executive functions: effects of training on academic achievement. *Psychol. Res.* doi: 10.1007/s00426-013-0537-1. [Epub ahead of print].

Tremblay, P., and Gracco, V. L. (2006). Contribution of the frontal lobe to externally and internally specified verbal responses: fMRI evidence. *Neuroimage* 33, 947–957. doi: 10.1016/j.neuroimage.2006.07.041

Troyer, A. K., Moscovitch, M., and Winocur, G. (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology* 11, 138–146. doi: 10.1037/0894-4105.11.1.138

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., and Stuss, D. (1998). Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia* 36, 499–504. doi: 10.1016/S0028-3932(97)00152-8

Unsworth, N., Spillers, G. J., and Brewer, G. A. (2010). Variation in verbal fluency: a latent variable analysis of clustering, switching, and overall performance. *Q. J. Exp. Psychol.* 64, 447–466. doi: 10.1080/17470218.2010.505292

Vandewater, E. A., Rideout, V. J., Wartella, E. A., Huang, X., and Shim, M. (2007). Digital childhood: electronic media and technology use among infants, toddlers, and preschoolers. *Pediatrics* 119, 1006–1015. doi: 10.1542/peds.2006-1804

Vygotsky, L. S. (1967). Play and its role in the mental development of the child. *Soviet Psychology*. 7, 6–18.

Welsh, M. C., Pennington, B. F., and Groisser, D. B. (1991). A normative-developmental study of executive function: a window on prefrontal function in children. *Dev. Neuropsychol.* 7, 131–149. doi: 10.1080/87565649109540483

Zelazo, P. D., Carlson, S. M., and Kesek, A. (2008). "The development of executive function in childhood," in *Handbook of Developmental Cognitive Neuroscience, 2nd Edn.* eds C. Nelson and M. Luciana (Cambridge, MA: MIT Press), 553–574.

Zelazo, P. D., and Lyons, K. E. (2012). The potential benefits of mindfulness training in early childhood: a developmental social cognitive neuroscience perspective. *Child Dev. Perspect.* 6, 154–160. doi: 10.1111/j.1750-8606.2012.00241.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 February 2014; accepted: 27 May 2014; published online: 17 June 2014.
*Citation: Barker JE, Semenov AD, Michaelson L, Provan LS, Snyder HR and Munakata Y (2014) Less-structured time in children's daily lives predicts self-directed executive functioning. Front. Psychol. **5**:593. doi: 10.3389/fpsyg.2014.00593*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Barker, Semenov, Michaelson, Provan, Snyder and Munakata. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 05 February 2014
doi: 10.3389/fpsyg.2014.00011

Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks

Virginia C. Mueller Gathercole[1]*, Enlli M. Thomas[2], Ivan Kennedy[2], Cynog Prys[3], Nia Young[2], Nestor Viñas Guasch[4], Emily J. Roberts[5], Emma K. Hughes[5] and Leah Jones[5]

[1] *Linguistics Program, English Department, Florida International University, Miami, FL, USA*
[2] *School of Education, Bangor University, Bangor, UK*
[3] *School of Social Sciences, Bangor University, Bangor, UK*
[4] *State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong, China*
[5] *School of Psychology, Bangor University, Bangor, UK*

Edited by:
Nicolas Chevalier, University of Colorado Boulder, USA

Reviewed by:
Ray Klein, Dalhousie University, Canada
Kenneth Paap, San Francisco State University, USA

****Correspondence:***
Virginia C. Mueller Gathercole, Linguistics Program, English Department, Florida International University, Modesto Maidique Campus, Miami, FL 33199, USA
e-mail: vmueller@fiu.edu

This study explores the extent to which a bilingual advantage can be observed for three tasks in an established population of fully fluent bilinguals from childhood through adulthood. Welsh-English simultaneous and early sequential bilinguals, as well as English monolinguals, aged 3 years through older adults, were tested on three sets of cognitive and executive function tasks. Bilinguals were Welsh-dominant, balanced, or English-dominant, with only Welsh, Welsh and English, or only English at home. Card sorting, Simon, and a metalinguistic judgment task (650, 557, and 354 participants, respectively) reveal little support for a bilingual advantage, either in relation to control or globally. Primarily there is no difference in performance across groups, but there is occasionally better performance by monolinguals or persons dominant in the language being tested, and in one case-in one condition and in one age group-lower performance by the monolinguals. The lack of evidence for a bilingual advantage in these simultaneous and early sequential bilinguals suggests the need for much closer scrutiny of what type of bilingual might demonstrate the reported effects, under what conditions, and why.

Keywords: executive function, bilingual children, language balance, language dominance, dimensional change card sort task, Simon task, metalinguistic task, Welsh bilinguals

INTRODUCTION

The question of bilinguals' linguistic and cognitive abilities relative to those of their monolingual counterparts has been the subject of intense study and scrutiny over the last century. Debates have examined children's and adults' capacities in a number of linguistic and cognitive realms. Early studies had mixed results concerning whether bilingualism was seen to have negative or positive effects on cognition, but many studies were flawed in that they did not control, e.g., for socioeconomic or cultural differences (Hakuta, 1986; Cummins, 1992; Oller and Pearson, 2002; Genesee et al., 2004). Recently, more controlled studies have indicated a complex picture. It is clear that in some ways, bilinguals' knowledge of certain aspects of their languages—in particular in lexical, morphological and syntactic realms—is affected by amount of exposure, so their abilities may show initial delays relative to those of their monolingual cohorts (Ben-Zeev, 1977; Umbel et al., 1992; Pearson et al., 1993, 1995; Pearson and Fernández, 1994; Gathercole, 2002a,b,c, 2007a,b; Gathercole and Hoff, 2007; Thomas and Gathercole, 2007; Gathercole and Thomas, 2009).

At the same time, bilinguals have been reported to show an advantage over their monolingual peers in the realms of metalinguistic abilities (Bialystok, 1993) and cognitive abilities related to executive function (Zelazo and Müller, 2002; Blair et al., 2005), involving selective attention, inhibition of attention, and switching attention in tasks with competing and misleading cues (Johnson, 1991; Bialystok et al., 2004; Hernandez Pardo et al., 2008). In these tasks, a high degree of cognitive control (Bialystok and Ryan, 1985) must be maintained, whether to inhibit irrelevant cues or to "detach" the verbal message from its reference (e.g., separate the linguistic form from its meaning). Successful completion entails ignoring conflicting or extraneous information. Bialystok (1993, 1999, 2001) argues that bilinguals have an advantage here because from the beginning of their use of two languages, bilinguals must control and suppress the use of one language while using the other (see also Cummins, 1976; Hakuta, 1986; Johnson, 1991; Green, 1998). This is purported to lead to more fully developed neurological mechanisms for controlling such attention, referred to as "executive function," which is relevant to the types of non-linguistic tasks mentioned (Bialystok and Ryan, 1985; Zelazo and Müller, 2002; Blair et al., 2005).

The advantage of bilinguals is reported, e.g., for the Stroop (1935) task, in which individuals are shown a color word written in a font of a color different from the color named by the word (e.g., *green* written in a red font) and are asked to name the color of the font, not read the word. In one study, Bialystok et al. (2008)

found that younger and older adults showed a greater Stroop effect (i.e., a greater cost in this condition than in non-conflict conditions) among monolinguals than bilinguals. They reasoned that monolinguals may show a greater Stroop effect because of their greater automaticity of reading. But even the performance of a group of monolinguals who were slower readers showed a greater Stroop effect than a group of bilinguals who were fast readers. (The important role that language and literacy abilities play in monolinguals' and bilinguals' performance is discussed further below).

In another task, the Simon task (Martin and Bialystok, 2003; Bialystok et al., 2004), participants are shown colored stimuli on the left or right side of a computer, and they are asked to press a key—one on the left or one on the right—according to the color of the stimulus on the computer. "Incongruent" trials, in which the stimulus and the correct key are on opposite sides of the computer, take more time (the "Simon effect") than "congruent" trials, in which the stimulus and the key are on the same side. Bilingual 4-year-olds show less of a Simon effect, and indeed also an advantage in the "congruent" cases, than monolinguals. (See below regarding the possibility of a global advantage in bilinguals.)

A third type of task is the "dimensional change card sort task" (Frye et al., 1995; Zelazo et al., 1996; Bialystok, 1999). In this task, participants are shown two target cards, one representing, e.g., a circle of one color (blue) and the other a square of another color (red), and then several other cards also showing opposite-colored circles and squares. The child is asked first to sort the items according to one dimension (e.g., color), and then to sort according to the other (shape). Bilingual children respond more accurately and more quickly than monolinguals on the second sort (Bialystok, 1999).

The bilingual advantage in control tasks is argued to also lead to superior performance by bilinguals in certain conditions of yet another type of task, a metalinguistic judgment task. Work conducted by Bialystok and colleagues (Bialystok, 1986, 1988; Barac and Bialystok, 2012) has argued that, while general performance on grammaticality judgment tasks, especially in judging ungrammatical sentences (e.g., "Why the dog is barking so loudly?"), seems to be related to level of knowledge of the language, there is one condition in which bilinguals are said to outperform monolinguals across the board, regardless of their level of bilingualism. That is on grammatically correct, but anomalous sentences, such as "Why is the cat barking so loudly?" (Bialystok, 1988: p. 565). The superior performance by bilinguals in this condition is attributed directly to superior executive control of bilinguals: This is "because attention normally directed to the meaning of the sentence [has] to be intentionally suppressed. Thus, the judgment require[s] high levels of control" (Bialystok, 1988, p. 565). Thus, "[o]n these problems, bilingual children consistently outperform monolingual children (Bialystok, 1986; Cromdal, 1999)" (Hermanto et al., 2012, p. 133).

Despite the many studies documenting such a cognitive advantage in bilinguals, some research has challenged the generality of the effect. Some have questioned the source of the effect, some have argued for better control over the choice of bilingual participants (e.g., Namazi and Thordardottir, 2010), and some have reported sporadic effects (Hilchey and Klein, 2011) or no bilingual effect (Paap and Greenberg, 2013), and there may be some with null effects that have not reached publication: Adesope et al. (2010) caution that there may be a "publication bias" against studies showing null or negative effects.

Yang and Lust (2004), for example, found no difference between monolingual and bilingual children's performance on a dimensional change card sort task but an advantage of bilinguals in an attentional network test, including a flanker task. They note that their monolinguals performed better on a vocabulary task, and so they suggest that language ability may have contributed to the lack of an effect for the card sort tasks; furthermore, their study controlled for the L1 languages of their participants, whereas many such studies pool participants from a variety of linguistic backgrounds and levels of proficiency. Variations in the first language backgrounds could have an effect on performance: Yang and Lust (2007) reported that children learning Korean and Chinese showed better performance on executive function tasks than those learning Spanish, regardless of linguality status (monolingual vs. bilingual), and in a systematic review of the literature, Adesope et al. (2010) reported significant differences in performance across distinct geographical and language groups, especially in relation to metalinguistic abilities.

Rosselli et al. (2002) also controlled for language background in a study of Spanish-English bilinguals' and monolinguals' performance on Stroop tasks. They found that bilinguals' performance was on the whole equivalent to monolinguals'. The one exception was that when asked to respond in English, bilinguals were generally slower than monolinguals, and Spanish-dominant bilinguals were slower than both English-dominant and balanced bilinguals. They suggest that the color naming effects may be related to vocabulary size. (See also Sumiya and Healy, 2004).

Similarly, in Chen and Ho (1986), Chinese L1-English L2 speakers in grade 2 through college performed Stroop tasks in Chinese and English; in some cases the language of the stimulus was the same as the language of the response, and in some different. The general finding was that within-language responding created greater interference than between-language responding, except for the youngest children. For these children, responses in English took longer with Chinese stimuli than with English stimuli. Since the younger children were less proficient in English, these results suggest that proficiency plays a role in the presence of the Stroop effect: the greater the proficiency, the more likely the within-language interference. Paap and Greenberg (2013) similarly report a lack of a bilingual advantage on a series of tasks when only highly proficient bilinguals are compared with monolinguals.

Socio-economic level might also contribute to results (Morton and Harper, 2007); monolingual and bilingual populations tested in some studies may have come from distinct socio-economic backgrounds (e.g., monolinguals from the general local population, bilinguals from L2 immigrants seeking higher education or from high SES academic parents choosing bilingual education for their children), and the effects of bilingualism may be more pronounced at some SES levels than at others (Woodard and Rodman, 2007). Hilchey and Klein (2011) point out the vast differences in sociocultural backgrounds of the bilingual vs.

monolinguals in a series of studies, and caution that there may be many such "hidden factors" other than linguality *per se* that lead to differences in performance. This point regarding SES is important because of recent work (Neville, 2009) indicating profound cognitive and neurological effects of SES level on attention in children. In a recent study, Paap and Greenberg (2013) tested monolingual and bilingual college students in California on a Simon task and a flanker task, and controlled for parental education. They found no significant difference between monolingual and bilinguals on either task. In another study (Duñabeitia et al., 2013), monolingual and bilingual children in the Basque country were carefully matched on a variety of skills (reading, arithmetic, verbal, IQ, etc.), and were tested for performance on a classic verbal Stroop task and a numerical Stroop task. These researchers consistently failed to find any significant difference in performance between the monolinguals and bilinguals.

Thus, the source and generality of experimental effects in bilinguals vs. monolinguals is not always clear, demonstrating the need for more well-controlled studies. Hilchey and Klein (2011) suggest:

> When these factors are not well controlled, a primary concern is that some of them might contribute or lead directly to what would appear to be bilingual processing advantages, and indeed, concerns of this sort have permeated the bilingualism literature. (p. 642).

The contributions of degree of proficiency in the language, SES factors, general cognitive abilities, age, and gender (and interactions between these) are still little understood in relation to bilinguals' and monolinguals' performance. Even the role of language dominance in the bilingual's performance is still unclear—it is not known to what extent various levels of language dominance might affect the cognitive benefits of bilingualism (Bialystok, 1988; Bialystok et al., 2004).

Furthermore, some have argued for a general cognitive advantage in bilinguals, not an advantage for inhibitory control (Hilchey and Klein, 2011). Hilchey and Klein (2011) review the evidence to date for a bilingual inhibitory control advantage (BICA), and conclude that there is little support for this position. In contrast, they argue, the evidence supports a more global bilingual executive processing advantage (BEPA) that leads to superior performance not only in conflict conditions (incongruent trials) but also in non-conflict conditions (congruent trials), particularly for RTs. They propose an alternative account, drawing on a conflict-monitoring system, to explain this global advantage; a similar account has been proposed by Costa et al. (2009). Paap and Greenberg (2013), like Hilchey and Klein (2011) failed to find in a series of studies any Group × Condition interactions revealing superior performance of bilinguals on conflict conditions. However, in contrast to Hilchey and Klein, Paap and Greenberg report no global advantages for bilinguals on their tasks. In fact, they argue, it is important that there is not a consistent pattern of performance by individuals across tasks: The failure to find consistent bilingual advantages across distinct components of executive processing challenges any theory of a unified account for results, even when a bilingual advantage is observed.

Adesope et al. (2010) note that often studies do not give clear information on the type of bilingual tested. In many studies, bilinguals are chosen as "balanced" on the basis of the fact that they have spoken both of their languages on a daily basis throughout their lives, but bilinguals lie on continua of dominance (Hakuta, 1987). Balanced bilinguals are not necessarily the same as those who use both of their languages on a daily basis (Grosjean, 1994; Grosjean and Li, 2003), and fully balanced bilinguals are quite rare (Hakuta, 1987).

Ultimately, the extent to which each factor contributes to performance is not well-understood. As Hilchey and Klein (2011, p. 643) say, "The onus is now on current investigative work to ensure that these factors are not influencing experimental outcomes."

The goal of the present study was to test performance on a series of executive function tasks in a carefully controlled study on bilinguals and monolinguals who grew up in the same context. The data come from Welsh-English bilinguals living in North West Wales. This group can provide insight into the effects of bilingualism in individuals who grow up as bilinguals—either as 2L1 simultaneous bilinguals or as early sequential bilinguals who begin the second language by age 4 at the latest—in comparison with monolinguals who are from the same sociocultural background.

In this study, we strictly divide the bilingual participants, first, according to the languages that their parents speak to them in the home—only Welsh at home (OWH), Welsh and English (WEH), or only English (OEH). In our work on children's acquisition of Welsh (Gathercole et al., 2001, 2005; Gathercole and Thomas, 2005; Thomas and Gathercole, 2005; Thomas et al., 2013) and on bilingual language transmission in Welsh homes (Gathercole, 2007a), we have found consistent differences across groups in the timing of acquisition or specific abilities in Welsh vs. English. The greater the exposure to Welsh, the earlier the child develops Welsh structures and vocabulary; the greater the exposure to English, the earlier the development of English forms; children who have equal exposure fall between these two groups (see Gathercole and Hoff, 2007; Gathercole, 2010; Thomas and Mayr, 2010).

The determination of relative "dominance" (where dominance is defined according to relative abilities in the two languages) across the three home language groups is not unproblematic, however. Typically, at initial stages, OWH children can be considered the most Welsh-dominant of the three types, WEH the most balanced, and OEH the most English-dominant. By the teen years, the differences across the groups become indistinguishable in English, but the OWH group still surpasses the others in Welsh. So OWH speakers may be considered the most balanced at older ages (see Gathercole et al., 2013, 2014).

In a previous study (Gathercole et al., 2010), we administered tapping tasks and a Stroop task to primary school aged children and teenagers. We examined the contributions of home language, language abilities and usage, general cognitive performance, and socioeconomic level to children's performance on these two tasks. Results revealed a complex picture of their contributions. In the case of the tapping task, in which there was a copy condition and a switch condition, the analyses showed an overall advantage at primary age in the OWH and OEH children, with monolingual English ("MonE") children performing least well, but there was no evidence of an advantage of any group in just a switch

task, or on difference scores. By teen age, the OWH and WEH children showed better performance than the MonE and OEH children. The follow-up analyses indicated, further, a high degree of association between tapping performance and general number abilities and pattern discrimination abilities, and supported the initial results showing superior performance among those bilinguals who began Welsh earlier and English later and who speak a high percentage of Welsh.

In the Stroop task, participants were tested either in Welsh or in English on four conditions, one of which was the classic Stroop condition. Analyses showed no home language effect in Welsh at either age. For English, by the teen years, there was no home language effect, including no difference between the monolinguals and the bilinguals. At the younger age, the WEH children showed an advantage over the OWH and MonE participants, but the OWH children showed inferior performance on a control condition in which they had to retrieve the color name from their lexical store. This supports the position of a bilingual advantage, here in the WEH children, but also important contributions of automaticity related to literacy. Follow-up analyses also confirmed important contributions to performance of balanced use of the two languages, of SES, of overall cognitive abilities, and of general linguistic knowledge, as measured by vocabulary scores.

In order to document more fully where and when a bilingual advantage might occur in this type of bilingual population—fully bilingual participants who grew up or are growing up as simultaneous bilinguals or early sequential bilinguals, we administered a series of tasks on Welsh-English bilinguals from seven age groups, across the lifespan. The following experiments report on card sorting tasks, Simon tasks, and a metalinguistic judgment task, providing further evidence on the influence of bilingualism on tasks related to executive function.

GENERAL RESEARCH METHOD

Participants in seven age groups (from 3 years of age through over 60 years of age) were administered several executive function tasks, including the card sorting, Simon, and metalinguistic tasks to be reported here. Participants were also administered, when possible, vocabulary tests in English (Dunn et al., 1982) and Welsh, tests of receptive grammatical knowledge in Welsh and English, and tests of general (non-executive function) cognitive abilities (McCarthy, 1972; Raven et al., 1983). Parents or the participants themselves filled out an extensive background questionnaire that included information on language use in the home and at school, parental language background, and parental education and professions. (We will report on effects involving non-language factors in a later study).

We predicted that the overall findings would be consistent with superior performance by bilinguals, especially the balanced bilinguals, over monolinguals. In the case of the card sorting tasks, the prediction was that this advantage would be observable in greater accuracy or faster reaction times of the (balanced?) bilinguals over the monolinguals in the tasks involving a switch of the parameters on which to base the sort; in the case of the Simon tasks, the prediction was that (balanced?) bilinguals would show an advantage in the conflict condition. For the metalinguistic task, the prediction was that (balanced?) bilinguals would show a particular advantage in the condition that involved grammatical but anomalous sentences.

PARTICIPANTS

A total of 650 children and adults participated in the card sort tasks, 557 in the Simon tasks, and 354 in the metalinguistic task. With the exception of the metalinguistic task (which was not administered to the preschoolers) participants took part in all studies. Differences in numbers are due to attrition. Participants were recruited through schools in and around North Wales, bilinguals from Gwynedd, Denbigh, and Conwy counties, and monolinguals from the Chester area, just across the Welsh border into England. Informed consent was obtained from participants or parents of participants. Across the tasks, participants fell into 7 age categories and four major home-language groups. Children came from 5 different age groups, around 3, 4, and 5 years of age, 8 years of age (henceforward "primary schoolers"), and 15 years of age (henceforward "teens" or "teenagers"); adults came from two groups, younger adults and older adults. (Exact ages will be reported with each task). All age groups performed (different versions of) a card sort task and a Simon task, and those of primary school age and above a metalinguistic task.

On the basis of the background questionnaires, participants were classified as either monolinguals ("MonE") or bilinguals coming from homes in which only Welsh was spoken ("OWH"), both Welsh and English were spoken ("WEH"), or only English was spoken ("OEH")[1].

CARD SORT

METHODS

Participants

The distribution of participants for the card sort tasks was as shown in **Table 1**. Mean ages are shown in Appendix A.

Stimuli and procedure

Three types of card sort task were given, according to participants' age groups. The school age children and adults were provided with a set of normal playing cards and were asked to sort them according to the experimenter's instructions. The specific sorting tasks differed, however, for the primary school age children from the older participants. The exact instructions and procedure for each are given in Appendix B. The youngest 3 age groups of children were given a simpler dimensional change card sort task, also described in Appendix B. In each case, participants were asked to sort the cards according to one criterion first, and then according to another criterion on a second (and in the case of older participants, additional) sort. Participants' accuracy and reaction times were recorded for every sort.

[1]A child was classified as OWH (OEH) if the parents reported at least 80% use of Welsh (English) in the home in speech to the child from birth to the present time, and adults were classified as OWH (OEH) according to the "origin home language," the home language patterns in their homes when they were children. Participants were classified as WEH if they received between 40 and 60% use of both languages in the home from their parents.

RESULTS

Cost

The "cost" associated with switching from one criterion to another in the sorting tasks was measured by the difference in performance between the first and second sorts (first minus second). Both the difference scores for accuracy and for reaction times were examined.

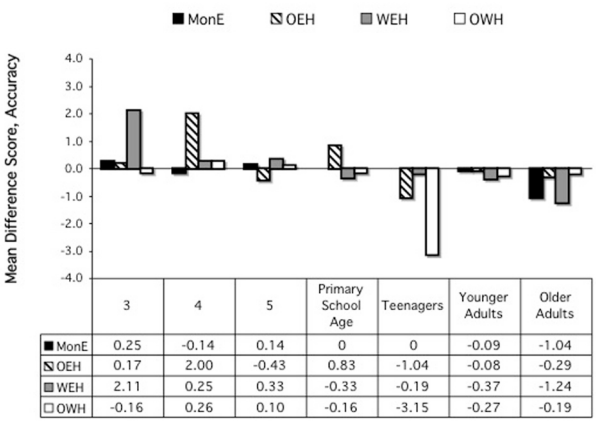
Accuracy

The difference scores for accuracy were entered into an ANOVA in which age and home language were entered as variables, with the difference score as the dependent measure. There was a main effect of age group, $F_{(6,\ 683)} = 4.83$, $p < 0.001$, with teens performing significantly better than the 3- and 4-year-olds, Scheffe's multiple comparisons, ps $= 0.002$.

There was also a significant interaction of Age Group × Home Language, $F_{(18,\ 683)} = 2.21$, $p = 0.003$. Performance is shown in **Figure 1**. Follow-up analyses at each age revealed that there was a difference by home language only for the teen group, $F_{(3,\ 105)} = 6.76$, $p < 0.001$. Scheffe's multiple comparisons showed that the OWH group outperformed (i.e., had less of a switch cost than) the MonE and the WEH groups, p's = 0.005, 0.004, respectively.

Table 1 | Participants, card sorting tasks.

Age group	MonE	OEH	WEH	OWH	TOT
3	14	20	16	21	71
4	29	19	18	18	84
5	19	21	21	19	80
Primary schoolers	25	22	20	29	96
Teens	20	24	26	34	104
Younger adults	28	20	28	30	106
Older adults	23	24	19	31	97
Total	161	151	150	188	650



	3	4	5	Primary School Age	Teenagers	Younger Adults	Older Adults
MonE	0.25	-0.14	0.14	0	0	-0.09	-1.04
OEH	0.17	2.00	-0.43	0.83	-1.04	-0.08	-0.29
WEH	2.11	0.25	0.33	-0.33	-0.19	-0.37	-1.24
OWH	-0.16	0.26	0.10	-0.16	-3.15	-0.27	-0.19

FIGURE 1 | Differences scores, accuracy, on card sort tasks, by age and home language.

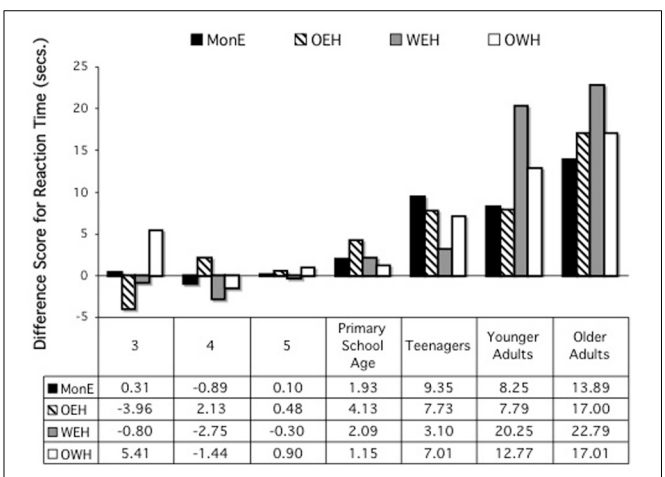
Reaction times

A second ANOVA examined the difference scores for reaction times. Again, age group and home language were entered as variables. This analysis revealed only a significant main effect of age group, $F_{(6,\ 674)} = 25.66$, $p < 0.001$. Scheffe's multiple comparisons revealed primarily differences between the two adult groups and the children's groups, with the older adults differing from all the children groups, all p's < 0.001, and the younger adults from all children groups except the teens, all p's < 0.001. The teens also differed significantly from the 4-year-olds, $p = 0.033$. There were no differences by home language (See **Figure 2**).

Global advantage?

To check for a possible global (BEPA) advantage for bilinguals, the data were reanalyzed, with separate tests conducted on the scores for accuracy and RTs on the first vs. second sorts by each age group. For accuracy, the only significant group effects were at the teen age group, which showed a significant HL effect, $F_{(3,\ 102)} = 5.94$, $p = 0.001$, and an interaction of HL × Sort, $F_{(3,\ 102)} = 6.76$, $p < 0.001$. These effects were due to the OWH group performing worse (at 46.6 correct) than all others (at 50.24–51.75 correct) on the first sort, $F_{(3,\ 120)} = 11.15$, $p < 0.001$, Scheffe's multiple comparisons, ps < 0.002.

For RTs, there were significant effects of HL group at ages 3 [$F_{(3,\ 71)} = 3.12$, $p = 0.031$], 4 [$F_{(3,\ 83)} = 5.43$, $p = 0.002$], 5 [$F_{(3,\ 79)} = 2.95$, $p = 0.038$], the teens [$F_{(3,\ 102)} = 12.46$, $p < 0.001$], and the younger adults [$F_{(3,\ 115)} = 4.61$, $p = 0.004$]. For the younger adults, there was also an interaction of HL × Sort, $F_{(3,\ 115)} = 5.60$, $p = 0.001$. In every case except for the teens, the Mons or English-dominant bilinguals were faster than one or more groups of the more balanced or Welsh-dominant bilinguals: at 3, Mon (34.91) < OWH (63.45), $p = 0.057$; at 4, Mon (21.68) < WEH (32.98), OWH (32.29), ps $= 0.009, 0.022$, respectively; at 5, OEH (18.73) tended to be faster than WEH (24.88), $p = 0.062$; at adults, Mon (45.07) < OWH (60.45), $p = 0.007$, Scheffe's multiple comparisons. For the teens, the Mon group was



	3	4	5	Primary School Age	Teenagers	Younger Adults	Older Adults
MonE	0.31	-0.89	0.10	1.93	9.35	8.25	13.89
OEH	-3.96	2.13	0.48	4.13	7.73	7.79	17.00
WEH	-0.80	-2.75	-0.30	2.09	3.10	20.25	22.79
OWH	5.41	-1.44	0.90	1.15	7.01	12.77	17.01

FIGURE 2 | Differences scores, reaction times, on card sort tasks, by age and home language.

slower (62.51) than all the bilingual groups (32.83–43.50), *p*s < 0.005.

DISCUSSION, CARD SORT

The results for the difference scores on the card sort tasks reveal little support for a bilingual advantage in relation to control (BICA), either for accuracy or reaction times, in relation to the costs related to a switch in the criteria to be followed in sorting. There was a single case, for accuracy at the teen years, in which the OWH children performed better than the MonE and WEH children; for RTs, MonE outperformed one or more bilingual group at ages 3, 4, 5, and younger adults; only among the teens were the bilinguals faster than the MonE children.

Similarly, the results on the absolute scores for accuracy and RTs on the first vs. second sort fail to support a global (BEPA) bilingual advantage. There was no difference by group at primary school age or among older adults; for 3-, 4-, and 5-year-olds and younger adults, the MonE or OEH participants outperformed the WEH and/or OWH participants; and for teens, the OWH group had lower accuracy rates than everyone else, but for RTs, this is the one place in which bilinguals outperformed monolinguals.

SIMON TASK

Two versions of the Simon Task, first created by Simon and Wolf (1963), were used in this study. We created one version specifically for younger children, and another for use with older children and adults.

PARTICIPANTS

The participants for the Simon tasks were distributed as in **Table 2**. The mean ages are shown in Appendix A.

STIMULI

Adult version

The adult version of the task involved a blue and a red square, which appeared either on the right or the left side of the computer screen. The participant's task was to press the Q on the computer if the blue square appeared and a P if the red square appeared.

Child version

The child version of the task involved a rabbit and a pig, who appeared sitting on top of a rock either on the right or the left side of the computer screen. The child's task was to touch a "button" on a touch screen, to indicate whether the rabbit or the pig appeared. The "buttons" showed either the rabbit or the pig, and the rabbit button always appeared at the bottom left of the screen and the pig button always appeared at the bottom right of the screen.

PROCEDURE

Participants were told, both verbally and in writing on the screen, to respond as quickly as possible to indicate which item appeared. If the blue square/rabbit appeared, the Q or the button on the left was to be pressed, and if the red square/pig appeared, the P or the button on the right was to be pressed. Between trials a "+" appeared in the center of the screen. The target item appeared on the screen half of the time on the left, and half the time on the right: in "congruent" trials, the target item appeared on the same side of the screen as the key or button to be pressed; in "incongruent" trials, the item appeared on the side of the screen opposite to that on which the key or button to be pressed was located. Three practice trials were given first, and then the target trials.

School age children and adults received 48 trials, 24 congruent, and 24 incongruent, in random order. The younger children received 16 trials, 8 congruent, and 8 incongruent. Accuracy of responses and reaction times were recorded electronically.

RESULTS

Youngest ages

Accuracy. An ANOVA was conducted in which condition (congruent, incongruent), age group, and home language were entered as independent variables and number correct responses as the dependent variable. There were main effects of condition, $F_{(1,\ 196)} = 27.25$, $p < 0.000$, and of age group, $F_{(2,196)} = 41.27$, $p < 0.000$, and an interaction of Condition × Age Group, $F_{(2,\ 196)} = 9.29$, $p < 0.000$. Children were more accurate in the congruent condition, with a mean of 7.07 correct, than in the incongruent condition, 6.37 correct, and performance increased between age 3, on the one hand, and 4 and 5 on the other, *p*s < 0.000, with means of 5.53, 7.20, and 7.43 at ages 3, 4, and 5, respectively.

Performance by each group is shown to the left in **Figures 3**, **4**, showing the congruent and incongruent conditions, respectively.

Table 2 | Participants, Simon tasks.

Age group	MonE	OEH	WEH	OWH	TOT
3	11	20	17	22	70
4	29	9	13	16	67
5	20	16	19	16	71
Primary schoolers	13	20	17	14	64
Teens	20	28	31	35	114
Younger adults	20	19	23	23	85
Older adults	20	23	17	24	84
Total	134	136	137	150	557

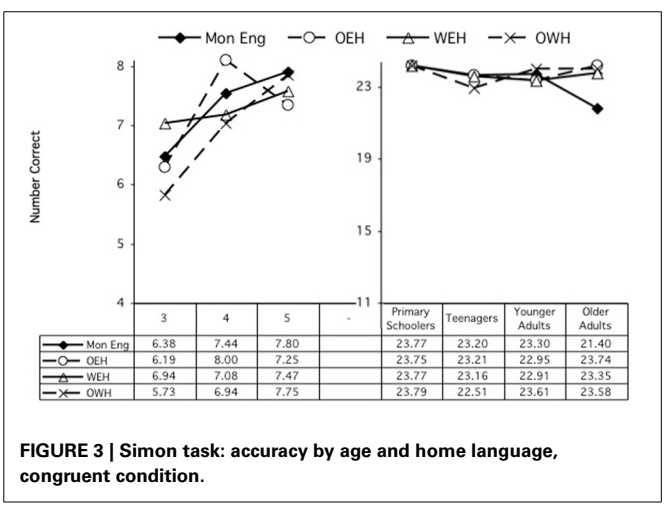


FIGURE 3 | Simon task: accuracy by age and home language, congruent condition.

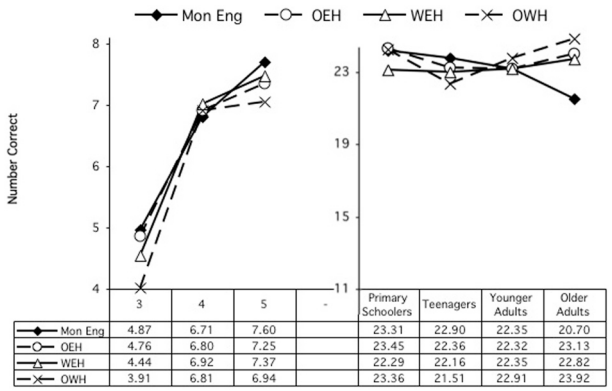


FIGURE 4 | Simon task: accuracy by age and home language, incongruent condition.

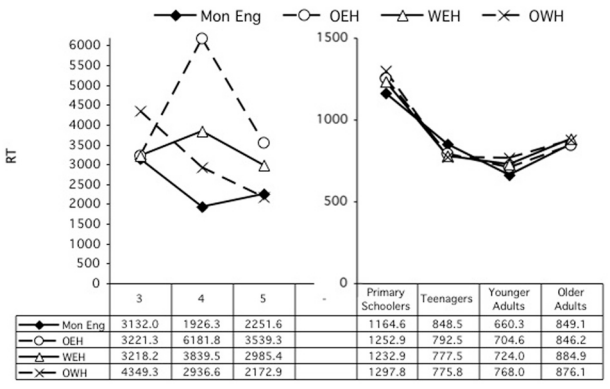


FIGURE 5 | Simon task: RT by age and home language, congruent condition.

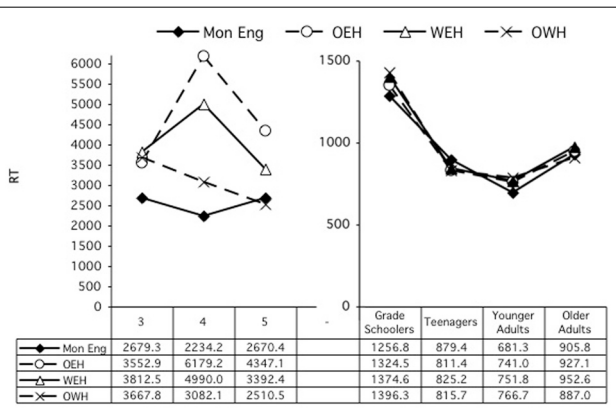


FIGURE 6 | Simon task: RT by age and home language, incongruent condition.

Follow up ANOVAs to examine the Condition × Age Group interaction looked at each age group separately. These revealed significant effects of condition (better on congruent) at ages 3 and 5, $F_{(1,\ 66)} = 22.65$, $p < 0.000$, and $F_{(1,\ 67)} = 6.88$, $p = 0.011$, respectively, but not at age 4.

There were no significant effects based on home language.

RTs. Similarly, an ANOVA was conducted involving the same independent variables to examine reaction time performance. This analysis revealed a main effect of condition, $F_{(1,\ 198)} = 4.02$, $p = 0.046$, and of home language, $F_{(3,\ 198)} = 3.41$, $p = 0.019$. The children were generally faster in the congruent condition, 3323.29 ms, than in the incongruent condition, 3590.3 ms. Mons were significantly faster overall, at 2482.1 ms, than OEH children, 4502.05 ms, $p = 0.002$, and nearly significantly than WEH children, 3707.3 ms, $p = 0.055$; OWH children were also significantly faster (3135.69 ms) than OEH children, $p = 0.040$. There were no other main or interaction effects. Performance in the congruent and incongruent conditions are shown to the left in **Figures 5**, **6**, which show the congruent and incongruent conditions, respectively.

School age children and above

Accuracy. An ANOVA was conducted in which condition (congruent, incongruent), age group, and home language were entered as independent variables and number correct as the dependent variable. There were main effects of condition, $F_{(1,\ 331)} = 39.81$, $p < 0.000$, and of age group, $F_{(3,\ 331)} = 2.73$, $p = 0.044$. Participants were generally more accurate in the congruent condition (23.25) than in the incongruent condition (22.58). And school-age children were more accurate than the other age groups (school-age: 23.44; teens: 22.63; younger adults: 22.84; older adults: 22.75), $ps < 0.05$.

There was also an interaction of Age Group × Home Language, $F_{(9,\ 331)} = 3.14$, $p = 0.001$. To explore this interaction, ANOVAs were conducted for each age group separately. Performance by each group is shown to the right in **Figures 3**, **4** (with the scale calibrated to show performance relative to that of the preschoolers). For the Primary Schoolers, there was no significant effect. For the teenagers, there were main effects of condition, $F_{(1,\ 110)} = 18.72$, $p < 0.000$, with better performance in the congruent condition (congruent: 23.02, incongruent: 22.23), and a trend in differences in performance by home language, $F_{(3,\ 110)} = 2.08$, $p = 0.107$. Pairwise comparisons revealed more accurate performance by the MonE participants (23.05) than the OWH participants (22.01), p =0.027 (with OEH and WEH in between, at 22.79 and 22.66 correct, respectively). For the younger adults, there was a main effect of condition, $F_{(1,\ 81)} = 19.63$, $p < 0.000$ (congruent: 23.29, incongruent: 22.48), but no effects involving home language. For the older adults, there were significant main effects of condition, $F_{(1,\ 80)} = 23.72$, $p < 0.000$, and of home language, $F_{(3,\ 80)} = 6.12$, $p = 0.001$. There was better performance on the congruent (23.02) than on the incongruent condition (22.49), and MonE participants (21.05) performed less well than all other groups, $ps = 0.004$ (OEH: 23.44, WEH: 23.09, OWH: 23.44).

Reaction times. Similarly, an ANOVA was conducted involving the same independent variables to examine reaction time performance. This analysis revealed a main effect of condition, $F_{(1,\ 330)} = 64.98$, $p < 0.001$, and of age group, $F_{(3,\ 330)} =$

144.42, $p < 0.001$. Participants were generally faster in the congruent condition, 903.481 ms, than in the incongruent condition, 956.080 ms. And all age groups had significantly different reaction times, all ps = 0.003, with the school age children the slowest (1287.55 ms), the young adults the fastest (724.72 ms), and the teens (815.75 ms) and older adults (891.11 ms) in between. There was also a significant interaction of Condition × Age Group, $F_{(3, 330)} = 6.22$, $p < 0.000$. There were no other main or interaction effects. Performance is shown to the right in **Figures 5**, **6** (with the scale adjusted in comparison with that for the preschoolers).

To explore the interaction of Condition × Age Group, separate ANOVAs were computed for each age group. Every age group showed faster performance on the congruent condition than on the incongruent condition: primary school age: $F_{(1, 59)} = 32.74$, $p < 0.000$; teens: $F_{(1, 110)} = 17.97$, $p < 0.000$; younger adults: $F_{(1, 81)} = 4.41$, $p = 0.039$; older adults: $F_{(1, 80)} = 9.90$, $p = 0.002$. The only group that showed an effect of home language was the younger adults, $F_{(3, 81)} = 3.47$, $p = 0.020$. In that group, the MonE participants were significantly faster overall (670.80) than both the WEH (737.94) and the OWH participants (767.35), ps = 0.031, 0.002, respectively.

SUMMARY, SIMON TASK

The results across the Simon tasks revealed that, consistent with predictions, all groups performed better on the congruent condition of the task than on the incongruent condition, but, inconsistent with predictions, there was little evidence of a bilingual advantage, either in accuracy of performance or in reaction times. Where there were effects involving home language, they were mixed. The MonE group often performed better or faster than one or more bilingual groups (for RTs in preschoolers and younger adults, for accuracy in teens); however, in one case the MonE group performed worse than the bilinguals (i.e., for accuracy among the older adults), and in another, the OWH participants patterned with the MonE participants in having faster RTs than OEH participants, in the preschool groups.

METALINGUISTIC TASK

METHOD

Participants

For the metalinguistic task, a total of 354 participants were tested, from four age groups: primary schoolers, teens, younger adults, and older adults. The distribution of participants by age group and home language was as shown in **Table 3**. The Monolingual English participants were given only the English task; all three bilingual home language groups were given both the English and the Welsh task.

The mean ages for each group are shown in Appendix A.

Table 3 | Participants, metalinguistic task, English and Welsh.

Age group	MonE	OEH	WEH	OWH	TOT
Primary schoolers	12	19	15	21	67
Teens	21	23	24	35	103
Younger adults	27	21	24	25	97
Older adults	22	21	17	27	87
Total	82	84	80	108	354

Stimuli

For both languages, 24 sentences were drawn up. In these, 6 types of structures were manipulated, and each type of structure was used in a grammatical meaningful sentence ("GM"), a grammatical, but anomalous sentence ("Gm"), an ungrammatical meaningful sentence ("gM"), and an ungrammatical anomalous sentence ("gm"). The 6 types of structures involved subject-verb agreement, irregular past tense formation, position of object pronouns, subject-auxiliary inversion in *wh-* questions, co-occurrence restrictions between the comparative form and the standard marker (*than*), and sequence of tenses. This design yielded 6 trials for each of the sentential conditions, GM, Gm, gM, and gm.

Examples of the English and Welsh sentences involving subject-verb agreement are shown in (1) in **Table 4**, and involving irregular past tense are shown in (2) in **Table 4**.

For the two languages, two versions of the sentences were drawn up. In the two versions, items that were grammatical and/or meaningful in one appeared as ungrammatical and/or anomalous in the other. For example, in one version, "Jim did his painting, so he bringed his brush to his dad to clean" occurred as gM, and in the other "Jim did his painting, so he bringed his brush to his dad to wear" occurred as gm. Bilingual participants heard the English sentences from one of these versions and Welsh sentences from the other. Monolinguals heard the English sentences from only one of the versions. The use of the two versions in each language across the participants was balanced.

Procedure

Participants heard sentences read to them orally. They were asked to judge whether a sentence was grammatical, and to correct it if it was ungrammatical. (See Appendix C for more details.) Participants were given 5 practice sentences, and then the target trials. The trial sentences were given in random order.

RESULTS

English

An ANOVA was conducted in which condition (GM, Gm, gM, gm), age, and home language were entered as independent variables and number correct responses as the dependent variable. There were significant main effects for all variables: condition, $F_{(3, 1014)} = 128.81$, $p < 0.000$; age, $F_{(3, 338)} = 63.55$, $p < 0.000$; home language, $F_{(3, 338)} = 2.99$, $p = 0.031$. Overall, participants performed differently on all conditions, pairwise comparisons, ps = 0.001, with performance best for GM (5.60 correct), next best for Gm (5.19), next for gM (4.88), and least good for gm (3.95). Similarly, all age groups performed significantly differently, all ps = 0.015, with improvement with age: primary schoolers: 3.79, teens: 4.92, younger adults: 5.30, older adults: 5.61. The effect of home language was due to significantly better performance overall by the OEH participants (5.11) over the WEH (4.73) and OWH (4.84) participants, pairwise ps = 0.005, 0.027, respectively. (MonE fell between the two extremes: 4.93).

Table 4 | Sample sentential stimuli for the metalinguistic task.

	GM	Gm	gM	gm
(1) SUBJECT-VERB AGREEMENT				
E	Today, Tommy is travelling to the zoo to visit the animals.	Today, Jenny is walking to the park to play with clouds.	Today, Billy am riding to the airport to see the planes.	Today, Mary are going to the supermarket to buy three pilots.
W	Heddiw, mae Tomi yn teithio i'r sw i weld yr anifeiliaid.	Heddiw, mae Jini yn cerdded i'r parc i chwarae ar y cymylau.	Heddiw, rydw Bili yn rhedeg i'r maes awyr i weld yr awyrennau. [requires *mae*, not *rydw*]	Heddiw, roeddwn Mari yn mynd i'r siop i brynnu tri peilot. [requires *roedd*, not *roeddwn*]
(2) IRREGULAR PAST TENSE				
E	Sam finished his work, so he gave his paper to the teacher to mark.	Sue ate her lunch, so she left her plate for the cook to break.	Jim did his painting, so he bringed his brush to his dad to clean.	Jan read the story, so she taked the book to the librarian to chew.
W	Gorffennod Sam ei waith, felly mi roddodd ei bapur i'r athrawes i'w farcio.	Bwytaodd Siwan ei chinio, felly gadawodd ei phlat i'r ddynes cinio dorri.	Peintiodd Jim ei lun, felly daethodd â'r brws i'w dad i olchi. [requires *daeth*, not *daethodd*]	Darllennodd Sian y stori, felly aethwyd hi â'r llyfr i'r llyfrgellydd i gnoi. [requires *aeth*, not *aethwyd*]

These main effects were modified by a significant interaction of Condition × Age, $F_{(9,\ 1014)} = 8.31$, $p < 0.000$, a near-significant effect of Condition × Home Language, $F_{(9,\ 1014)} = 1.84$, $p = 0.058$, and a significant interaction of Condition × Age × Home Language, $F_{(27,\ 1014)} = 1.55$, $p = 0.036$. To examine the interactions, performance by each age group was analyzed separately.

Performance at each age is shown in **Figure 7**. The primary schoolers showed a significant main effect of condition, $F_{(3,\ 189)} = 40.75$, $p < 0.000$. Performance on all conditions was significantly different, ps < 0.000, except for the Gm and gM conditions, which reached near-significance, $p = 0.073$. There was a near-significant interaction of Condition × Home Language, $F_{(9,\ 189)} = 1.83$, $p = 0.066$. Follow-up analysis revealed a significant difference in performance on the gM sentences, $F_{(3,\ 63)} = 3.42$, $p = 0.022$, with OWH children performing lower than the OEH children, $p = 0.044$. The teens likewise showed a significant effect of condition, $F_{(3,\ 297)} = 38.52$, $p < 0.000$, with significant differences across all conditions, ps = 0.001, except for the Gm and gM conditions, $p = 0.238$. There were no other differences among the teens. The younger adults also showed an effect of condition, $F_{(3,\ 279)} = 24.48$, $p < 0.000$, with all conditions significantly different, ps = 0.036, except for the GM and Gm sentences, which were nearly significantly different, $p = 0.068$. There were no other differences among the younger adults. The older adults likewise showed a significant effect of condition, $F_{(3,\ 249)} = 21.29$, $p < 0.001$, but here performance differed only on the gm condition relative to all the others, ps < 0.000. There were no other differences.

Welsh

An ANOVA was similarly conducted examining performance on the Welsh sentences. There were significant main effects for all variables: condition, $F_{(3,\ 780)} = 169.56$, $p < 0.001$; age, $F_{(3,\ 260)} = 56.80$, $p < 0.001$; home language, $F_{(2,\ 260)} = 3.90$, $p = 0.021$. Overall, participants performed differently on all conditions, pairwise comparisons, ps $= 0.000$, with performance best for GM (5.56 correct), next best for Gm (5.10), next for gM (4.14), and least good for gm (3.52). Similarly, most age groups performed significantly differently, all ps $= 0.001$, except for the younger and older adults, who did not differ significantly, $p = 0.144$. Performance improved with age: primary schoolers: 3.32, teens: 4.59, younger adults: 5.09, older adults: 5.31. The effect of home language was due to significantly better performance overall by the OWH participants (4.79) and the WEH participants (4.50) over the OEH participants (4.45), pairwise ps $= 0.033$, 0.012, respectively.

These main effects were modified by a significant interaction of Condition × Age, $F_{(9,\ 780)} = 12.12$, $p < 0.000$, and of Condition × Home Language, $F_{(6,\ 780)} = 3.03$, $p = 0.006$. There were no other interactions. Follow-up analyses examined these interactions.

Performance at each age is shown in **Figure 8**. First, each age group was examined separately to explore the Condition × Age interaction. Analyses revealed that the Condition × Age interaction reflects the fact that performance differed on all conditions for the primary schoolers, pairwise ps $= 0.024$, but for the other age groups, all but the GM vs. Gm conditions differed, pairwise ps $= 0.002$. The Condition × Home Language interaction was explored by examining each condition separately. Analysis revealed that performance differed by home language only on the gM condition, $F_{(2,\ 269)} = 2.95$, $p = 0.054$. The OWH participants performed significantly better here (4.53) than the OEH participants (3.88), $p = 0.049$.

SUMMARY, METALINGUISTIC TASK

The results of the metalinguistic tasks also failed to reveal a bilingual advantage, either overall or in the crucial Gm condition, which requires the greatest levels of inhibitory control. This is contrary to expectations, according to the proposal of an executive function advantage by bilinguals in this condition. In

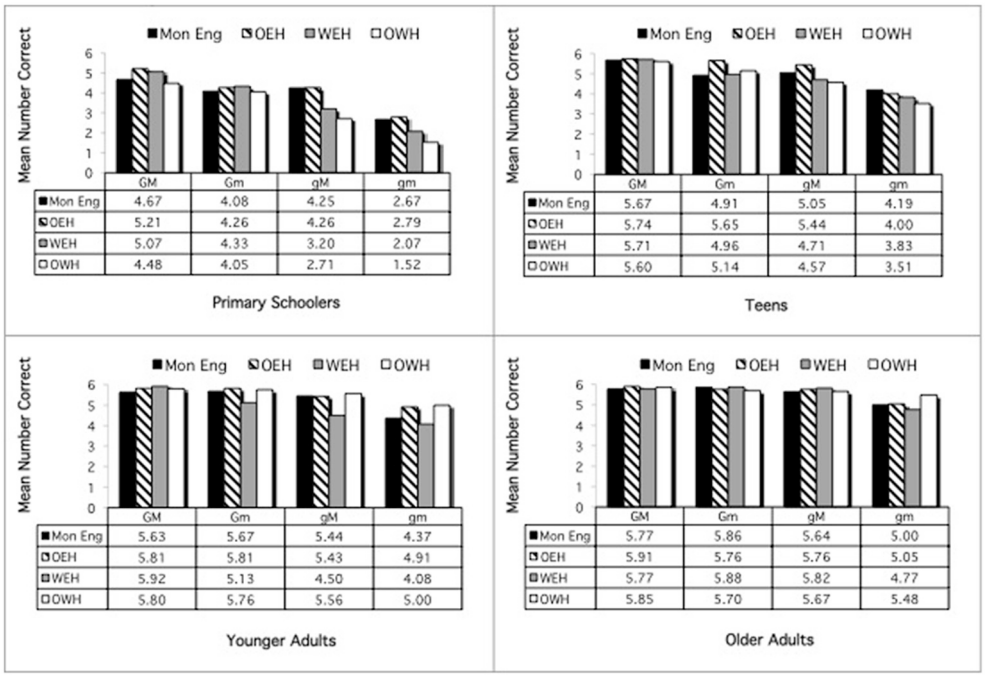


FIGURE 7 | Metalinguistic task: English.

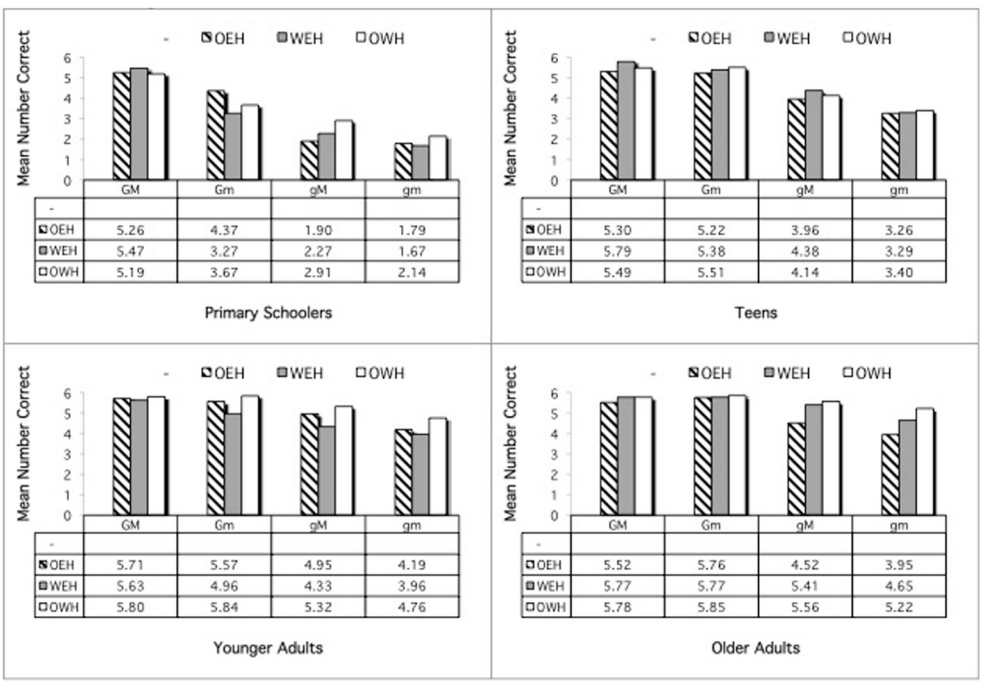


FIGURE 8 | Metalinguistic task: Welsh.

accordance with predictions related to language ability, in contrast, home language, when it mattered, showed an advantage in the direction of the bilingual group that was dominant in the given language. That is, for English, the OEH children performed the best of the bilinguals; and in the primary age group, on the gM condition (which requires greater levels of sentence analysis than control of attention), the OEH children outperformed the OWH children; in contrast in Welsh, the OWH and WEH

participants outperformed the OEH participants, and specifically in the gM condition, the OWH participants outperformed the OEH participants.

DISCUSSION

These experiments reveal that on three sets of executive function tasks, performance by this group of simultaneous and early sequential bilinguals fails to provide support for an overall bilingual advantage at any of the seven ages tested here. The card sorting tasks failed to show an overall advantage of bilinguals, either in relation to the "cost" of the switch or in relation to an overall performance advantage. On the Simon task, performance was generally similar across groups, or the monolinguals generally had the advantage; in many cases, the monolinguals (or in one case, the OEH bilinguals) were faster or more accurate than one or more groups of bilinguals. In one case, however, the OWH bilinguals were, like the monolinguals, also faster than the OEH and WEH bilinguals, and in one case, the monolinguals were less accurate than the bilinguals (at the older adult group). On the metalinguistic task, again where there were differences, the differences were in the direction of those dominant in the language being tested outperforming those who were less dominant, most importantly, even in the Gm condition, where executive control was predicted to favor bilinguals.

It should be noted that this evidence showing little support for the bilingual advantage was accompanied in every case by robust evidence supporting predictions not related to home language. For example, performance in congruent conditions was always superior to performance in incongruent conditions, both in accuracy and in RTs (similar to findings in Kousaie and Phillips, 2012; Paap and Greenberg, 2013; Duñabeitia et al., 2013); changes with age in children always showed better performance with age; changes with age in adults often showed decreased performance at the older ages; judgments of grammaticality were better with grammatical sentences than with ungrammatical sentences. This indicates that the tasks here elicited performance as predicted in all major ways except for one, in relation to bilingualism.

The absence of strong support for the position of a bilingual advantage on these executive tasks, as in our earlier work (and in some forthcoming work from Clare et al., submitted) is striking. This study examined a large number of fully fluent, simultaneous and early sequential bilinguals, homogeneous in cultural and educational backgrounds, and homogeneous with those of the monolinguals. While it is possible that language abilities contributed to performance on the card sorting and metalinguistic tasks, the Simon task is a classic task used to examine EF performance. The results here suggest that whatever mechanisms yield superior performance in other studies in relation to bilinguals and control may be less relevant to simultaneous and early sequential bilinguals.

As noted above, in many studies, the participants are L2 bilinguals (or not clearly defined, Adesope et al., 2010). The process of acquiring two languages and the relationship between the bilingual's two languages are clearly different in simultaneous bilinguals than in L2 bilinguals (see, e.g., Li, 2010), and one can predict that the use of language in the former group is likely to be more automatic and less effortful than in the latter group. This may make the theoretical issues surrounding control in bilinguals less relevant to simultaneous bilinguals than to L2 bilinguals. Paap and Greenberg (2013) point out that one of the background assumptions for theories of a bilingual advantage in EF is that "the amount of EP recruited by bilinguals during language comprehension and production is greater than that employed by monolinguals" (p. 255), but that speaking any language, whether bilingually or monolingually, involves a great deal of monitoring, switching, and inhibitory control. They add:

> To provide just a few examples, conversational participants must monitor the environment for signals regarding turn-taking, misunderstandings, possible use of sarcasm, changes of topic, or changes in register contingent upon who enters or leaves the conversation. These lead to switches from speaker to listener, switches from one knowledge domain to another, and so forth. Although monolinguals do not need to suppress translation equivalents during production, they incessantly make word choices among semantically and syntactically activated candidates that include synonyms, hypernyms, and hyponyms. In addition monolinguals must use context to suppress irrelevant meaning of homographs during comprehension (p. 256).

It is worth considering as well the extent to which the theory surrounding a bilingual advantage in relation to control hinges on a modular approach to language. If the two languages spoken by a bilingual are separate, then this would necessarily involve some mechanism for switching back and forth between the two languages. Consider, however, a less modular model of language. Under a computational model of language acquisition and language use, for example, the processes involved in language use can be seen more as involving activation of links than switches between two separate (but related) systems. The links within a language will be stronger than across languages, but both languages appear to be "on line" at all moments (see, e.g., Lam and Dijkstra, 2010). In fully fluent simultaneous bilinguals, in contrast to, e.g., recent or less fluent L2 language learners, the automaticity of their linguistic knowledge in both languages may mean that whatever "switching" they are carrying out is a function of the contexts of speech, just as it is for monolinguals. Less fluent bilinguals, L2 learners, on the other hand, may need to conduct a greater level of control in every linguistic choice they make. It is striking that much of the literature in which no bilingual advantage has been found has involved fully fluent bilingual communities such as this one in North Wales and the Basque Country (e.g., Duñabeitia et al., 2013).

These questions are deserving of much closer scrutiny in future research. The choice of participants in studies of this type needs to be controlled more carefully in the future, so that we can better define exactly who shows an advantage in performance, under what conditions, and why.

ACKNOWLEDGMENTS

This research was supported by ESRC grant RES-062-23-0175 on Cognitive effects of bilingualism across the lifespan and by ESRC/HEFCW grant RES-535-30-0061 for the ESRC Centre for Research on Bilingualism in Theory and Practice, which we gratefully acknowledge. We also wish to express our thanks to all of the schools, parents, children, and adults who participated in this work; without their very helpful cooperation this study could

never have taken place. Thank you also to project students and research assistants who helped in data collection: Carys Blunt, Angela H. Clifford, Lowri Cunnington, Nikki Davies, Sinead Dolan, Susan D. Elliott, Katie Gibbins, Bethan Gritten, Lowri Hadden, Josselyn Hellriegel, Catherine L. Hogan, Siwan Long, Jason H. McEvoy, Sarah N. Pearce, Hannah Perryman, Emily Roberts, Sioned Roberts, Kathryn Sharp, Heather Shawcross, and Lesley Waiting. Finally, thank you to the reviewers for their helpful and insightful suggestions for revision.

REFERENCES

Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803

Barac, R., and Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: role of language, cultural background, and education. *Child Dev.* 83, 413–422. doi: 10.1111/j.1467-8624.2011.01707.x

Ben-Zeev, S. (1977). "Mechanisms by which childhood bilingualism affects understanding of language and cognitive structures," in *Bilingualism: Psychological, Social, and Educational Implications,* ed P. Hornby (New York, NY: Academic Press), 29–55.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Dev. Psychol.* 24, 560–567. doi: 10.1037/0012-1649.24.4.560

Bialystok, E. (1993). "Metalinguistic awareness: the development of children's representations of language," in *Systems of Representation in Children: Development and Use*, eds C. Pratt and A. Garton (London: Wiley), 211–233.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* 70, 636–644. doi: 10.1111/1467-8624.00046

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition.* New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511605963

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859

Bialystok, E., and Ryan, E. B. (1985). Toward a definition of metalinguistic skill. *Merrill Palmer Q.* 31, 229–251.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Dev.* 57, 498–510. doi: 10.2307/1130604

Blair, C., Zelazo, P. D., and Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Dev. Neuropsychol.* 28, 561–571. doi: 10.1207/s15326942dn2802_1

Chen, H.-C., and Ho, C. (1986). Development of Stroop interference in Chinese-English bilinguals. *J. Exp. Psychol.* 12, 397–401. doi: 10.1037/0278-7393.12.3.397

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149. doi: 10.1016/j.cognition.2009.08.001

Cromdal, J. (1999). Childhood bilingualism and metalinguistic skills: analysis and control in young Swedish-English bilinguals. *Appl. Psycholinguist.* 20, 1–20. doi: 10.1017/S0142716499001010

Cummins, J. (1976). The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypotheses. *Work. Pap. Bilingualism* 9, 1–43.

Cummins, J. (1992). Bilingual education and English immersion: the Ramirez report in theoretical perspective. *Biling. Res. J.* 16, 91–104. doi: 10.1080/15235882.1992.10162630

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., et al. (2013). The inhibitory advantage in bilingual children revisited: myth or reality? *Exp. Psychol.* 1–18. doi: 10.1027/1618-3169/a000243. [Epub ahead of print].

Dunn, L. M., Dunn, L. M., and Whetton, C. (1982). *The British Picture Vocabulary Scale: Manual for the Short and Long Forms.* London: NFER-NELSON.

Frye, D., Zelazo, P. D., and Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cogn. Dev.* 10, 483–527. doi: 10.1016/0885-2014(95)90024-1

Gathercole, V. C. M. (2002a). "Command of the mass/count distinction in bilingual and monolingual children: an English morphosyntactic distinction," in *Language and Literacy in Bilingual Children,* eds D. K. Oller and R. E. Eilers (Clevedon: Multilingual Matters), 175–206.

Gathercole, V. C. M. (2002b). "Grammatical gender in bilingual and monolingual children: a Spanish morphosyntactic distinction," in *Language and Literacy in Bilingual Children,* eds D. K. Oller and R. E. Eilers (Clevedon: Multilingual Matters), 207–219.

Gathercole, V. C. M. (2002c). "Monolingual and bilingual acquisition: learning different treatments of *that*-trace phenomena in English and Spanish," in *Language and Literacy in Bilingual Children,* eds D. K. Oller and R. E. Eilers (Clevedon: Multilingual Matters), 220–254.

Gathercole, V. C. M. (ed.). (2007a). *Language Transmission in Bilingual Families in Wales.* Cardiff: Welsh Language Board.

Gathercole, V. C. M. (2007b). Miami and North Wales, so far and yet so near: constructivist account of morpho-syntactic development in bilingual children. *Int. J. Biling. Educ. Biling.* 10, 224–247. doi: 10.2167/beb442.0

Gathercole, V. C. M. (2010). Interface or face-to-face? The profiles and contours of bilinguals and specific language impairment. *Appl. Psycholinguist.* 31, 58–69. doi: 10.1017/S0142716409990439

Gathercole, V. C. M., and Hoff, E. (2007). "Input and the acquisition of language: three questions," in *The Handbook of Language Development,* eds E. Hoff and M. Shatz (New York, NY: Blackwell Publishers), 107–127.

Gathercole, V. C. M., Laporte, N., and Thomas, E. M. (2005). "Differentiation, carry-over, and the distributed characteristic in bilinguals: structural "mixing" of the two languages?," in *ISB4: Proceedings of the 4th International Symposium on Bilingualism,* eds J. Cohen, K. McAlister, K. Rolstad, and J. MacSwan (Somerville, MA: Cascadilla Press), 838–851.

Gathercole, V. C. M., Pérez Tattam, R., and Stadthagen-González, H. (2014). "Bilingual construction of two systems: to interact or not to interact?" in *Approaches to the Study of Bilingualism,* eds E. M. Thomas and I. Mennen (Bristol: Multilingual Matters), 63–89.

Gathercole, V. C. M., and Thomas, E. M. (2005). "Minority language survival: input factors influencing the acquisition of Welsh," in *ISB4: Proceedings of the 4th International Symposium on Bilingualism,* eds J. Cohen, K. McAlister, K. Rolstad, and J. MacSwan (Somerville, MA: Cascadilla Press), 852–874.

Gathercole, V. C. M., and Thomas, E. M. (2007). *Prawf Geirfa Cymraeg: Fersiwn 7-11. Welsh Vocabulary Test: Version 7-11*. Available online at: www.pgc.bangor.ac.uk

Gathercole, V. C. M., and Thomas, E. M. (2009). Bilingual first-language development: dominant language takeover, threatened minority language take-up. *Bilingualism* 12, 213–237. doi: 10.1017/S1366728909004015

Gathercole, V. C. M., Thomas, E. M., Jones, L., Viñas Guasch, N., Young, N., and Hughes, E. K. (2010). Cognitive effects of bilingualism: digging deeper for the contributions of language dominance, linguistic knowledge, socioeconomic status, and cognitive abilities. *Int. J. Biling. Educ. Biling.* 13, 617–665. doi: 10.1080/13670050.2010.488289

Gathercole, V. C. M., Thomas, E. M., and Laporte, N. (2001). The acquisition of grammatical gender in Welsh. *J. Celtic Lang. Learn.* 6, 53–87.

Gathercole, V. C. M., Thomas, E. M., Roberts, E. J., Hughes, C., and Hughes, E. K. (2013). "Why assessment needs to take exposure into account: vocabulary and grammatical abilities in bilingual children," in *Issues in the Assessment of Bilinguals*, ed V. C. M. Gathercole (Bristol: Multilingual Matters), 20–55.

Genesee, F., Paradis, J., and Crago, M. (2004). *Dual Language Development and Disorders: a Handbook on Bilingualism and Second Language Learning*. Baltimore, MD: Paul Brookes.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism* 1, 67–81. doi: 10.1017/S1366728998000133

Grosjean, F. (1994). "Individual bilingualism," in *The Encyclopedia of Language and Linguistics*, ed R. Asher (Oxford: Pergamon Press), 1656–1660.

Grosjean, F., and Li, P. (2003). Commentary, imaging bilinguals: when the neurosciences meet the language sciences. *Bilingualism* 6, 159–165. doi: 10.1017/S136672890300110X

Hakuta, K. (1986). *The Mirror of Language: the Debate on Bilingualism*. New York, NY: Basic Books.

Hakuta, K. (1987). Degree of bilingualism and cognitive ability in mainland Puerto Rican children. *Child Dev.* 58, 1372–1388. doi: 10.2307/1130628

Hermanto, N., Moreno, S., and Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: How bilingual? *Int. J. Biling. Educ. Biling.* 15, 131–145. doi: 10.1080/13670050.2011.652591

Hernandez Pardo, M., Costa, A., and Sebastián-Gallés, N. (2008). "The bilingual advantage in attentional control: monitor and conflict resolution processes," in *Paper presented at International Conference on Models of Interaction in Bilinguals* (Bangor).

Hilchey, M. D., and Klein, R. M. (2011). Are there bilingual advantages on non-linguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonom. Bull. Rev.* 18, 625–658. doi: 10.3758/s13423-011-0116-7

Johnson, J. (1991). "Cognitive processes in bilingualism and their cognitive growth effects," in *Language Processes in Bilingual Children,* ed E. Bialystok (Cambridge: Cambridge University Press), 193–221.

Kousaie, S., and Phillips, N. A. (2012). Conflict monitoring and resolution: are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Res.* 1446, 71–90. doi: 10.1016/j.brainres.2012.01.052

Lam, K. J. Y., and Dijkstra, T. (2010). Word repetition, masked orthographic priming, and language switching: bilingual studies and BIA+ simulations. *Int. J. Biling. Educ. Biling.* 13, 487–503. doi: 10.1080/13670050.2010.488283

Li, P. (2010). "Meaning in acquisition: semantic structure, lexical organization, and crosslinguistic variation," in *Routes to Language: Studies in Honor of Melissa Bowerman,* ed V. M. Gathercole (New York, NY: Psychology Press), 257–283.

Martin, M. M., and Bialystok, E. (2003). "The development of two kinds of inhibition in monolingual and bilingual children: Simon vs. Stroop," in *Poster Presented at the Meeting of The Cognitive Development Society* (Park City, Utah).

McCarthy, D. (1972). *McCarthy Scales of Children's Abilities*. New York, NY: The Psychological Corporation.

Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x

Namazi, M., and Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *Int. J. Biling. Educ. Biling.* 13, 597–616. doi: 10.1080/13670050.2010.488288

Neville, H. (2009). "Bilingualism: a window into neuroplasticity," in *Keynote Paper Presented at Conference on Neurobilingualism: Bilingual Functioning from Infancy to Adulthood* (Bangor University).

Oller, D. K., and Pearson, B. Z. (2002). "Assessing the effects of bilingualism: a background," in *Language and Literacy in Bilingual Children*, eds D. K. Oller and R. E. Eilers (Clevedon: Multilingual Matters), 3–21.

Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002

Pearson, B. Z., and Fernández, S. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Lang. Learn.* 44, 617–653. doi: 10.1111/j.1467-1770.1994.tb00633.x

Pearson, B. Z., Fernández, S., and Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: comparison to monolingual norms. *Lang. Learn.* 43, 93–120. doi: 10.1111/j.1467-1770.1993.tb00174.x

Pearson, B. Z., Fernández, S., and Oller, D. K. (1995). Cross-language synonyms in the lexicons of bilingual infants: one language or two? *J. Child Lang.* 22, 345–368. doi: 10.1017/S030500090000982X

Raven, J. C., Court, J. H., and Raven, J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 3) - Standard Progressive Matrices*. London: Lewis.

Rosselli, M., Ardila, A., Santisi, M. N., Arecco Model, R., Salvatierra, J., Conde, A., et al. (2002). Stroop effect in Spanish–English bilinguals. *J. Int. Neuropsychol. Soc.* 8, 819–827. doi: 10.1017/S1355617702860106

Simon, J. R., and Wolf, J. D. (1963). Choice reaction times as a function of angular stimulus-response correspondence and age. *Ergonomics* 6, 99–105.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Sumiya, H., and Healy, A. F. (2004). Phonology in the bilingual Stroop effect. *Mem. Cogn.* 32, 752–758. doi: 10.3758/BF03195865

Thomas, E. M., and Gathercole, V. C. M. (2005). "Obsolescence or survival for Welsh in the face of English dominance?" in *ISB4: Proceedings of the 4th International Symposium on Bilingualism,* eds J. Cohen, K. McAlister, K. Rolstad, and J. MacSwan (Somerville, MA: Cascadilla Press), 2233–2257.

Thomas, E. M., and Gathercole, V. C. M. (2007). Children's productive command of grammatical gender and mutation in Welsh: an alternative to rule-based learning. *First Lang.* 27, 251–278. doi: 10.1177/0142723707077056

Thomas, E. M., and Mayr, R. (2010). "Children's acquisition of Welsh in a bilingual setting: a psycholinguistic perspective," in *Welsh in the Twenty-First Century*, ed D. Morris (Cardiff: University of Wales Press), 99–117.

Thomas, E. M., Williams, N., Jones, L. l. A., Davies, S., and Binks, H. (2013). Acquiring complex structures under minority language conditions: bilingual acquisition of plural morphology in Welsh. *Bilingualism.* doi: 10.1017/S1366728913000497

Umbel, V. M., Pearson, B. Z., Fernández, M. C., and Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Dev.* 63, 1012–1020. doi: 10.2307/1131250

Woodard, E., and Rodman, A. (2007). "Socio-economic status as a mediating factor in the relationship between bilingualism and the cognitive development of pre-school children," in *Paper presented at the annual meeting of the National Communication Association 93rd Annual Convention,* (Chicago, IL).

Yang, S., and Lust, B. (2004). "Testing effects of bilingualism on executive attention: comparison of cognitive performance on two non-verbal tests," in *Paper presented at Boston University Conference on Language Development* (Boston: BUCLD), 29.

Yang, S., and Lust, B. (2007). "Cross-linguistic differences in cognitive effects due to bilingualism: experimental study of lexicon and executive attention in two typologically distinct language groups," in *Proceedings of Boston University Conference on Language Development (BUCLD) 31*, Vol. 2 eds Heather Caunt-Nulton, Samantha Kulatilake, and I-hao Woo (Somerville, MA: Cascadilla Press), 692–703.

Zelazo, P. D., Frye, D., and Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cogn. Dev.* 11, 37–63. doi: 10.1016/S0885-2014(96)90027-1

Zelazo, P. D. and Müller, U. (2002). "Executive functions in typical and atypical development," in *Handbook of Childhood Cognitive Development*, ed U. Goswami (Oxford: Blackwell), 445–469. doi: 10.1002/9780470996652.ch20

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 September 2013; accepted: 07 January 2014; published online: 05 February 2014.

Citation: Gathercole VCM, Thomas EM, Kennedy I, Prys C, Young N, Viñas Guasch N, Roberts EJ, Hughes EK and Jones L (2014) Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. Front. Psychol. ***5****:11. doi: 10.3389/fpsyg.2014.00011*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Gathercole, Thomas, Kennedy, Prys, Young, Viñas Guasch, Roberts, Hughes and Jones. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX A

MEAN AGES OF PARTICIPANTS

Card sort tasks

Mean ages by group were as follows:

3: mean: 3.0; range of means in the individual home language groups: 2.11–3.2 (individual range: 2.0–3.6)

4: mean: 4.2; range of means in the individual home language groups: 4.1–4.3 (individual range: 3.7–4.9)

5: mean: 5.5; range of means in the individual home language groups: 5.4–5.6 (individual range: 4.10–6.0)

Primary Schoolers: mean: 8.0; range of means in the individual home language groups: 7.8–8.5 (individual range: 7.0–8.11)

Teens: mean: 14.9, range of means in the individual home language groups: 14.7–14.11 (individual range: 13.0–16.0)

Younger Adults: mean: 24.8; range of means in the individual home language groups: 23.5–26.2 (individual range: 18.0–39.0)

Older Adults: mean: 67.4; range of means in the individual home language groups: 65.1–69.1 (individual range: 57.0–90.0)

Simon task

Mean ages for each group were as follows:

3: mean: 3.0; range of means in the individual home language groups: 2.11–3.3 (individual range: 2.0–3.6)

4: mean: 4.2; range of means in the individual home language groups: 4.1–4.5 (individual range: 3.5–4.11)

5: mean: 5.4; range of means in the individual home language groups: 5.4–5.6 (individual range: 4.10–6.0)

Primary School: mean: 8.2; range of means in the individual home language groups: 7.9–8.5 (individual range: 7.0–8.11)

Teens: mean: 14.9; range of means in the individual home language groups: 14.8–14.11 (individual range: 13.0–16.0)

Younger Adults: mean: 25.5; range of means in the individual home language groups: 23.4–26.6 (individual range: 18.3–39.0)

Older Adults: mean: 67.6; range of means in the individual home language groups: 66.2–68.4 (individual range: 57.6–90.0)

Metalinguistic task

The mean ages were as follows:

Primary School: mean: 8.2; range of means in the individual home language groups: 7.10–8.4 (individual range: 7.0–8.11)

Teens: mean: 14.10; range of means in the individual home language groups: 14.8–15.0 (individual range: 13.0–16.0)

Younger Adults: mean: 25.4; range of means in the individual home language groups: 23.2–27.2 (individual range: 18.3–39.0)

Older Adults: mean: 67.7; range of means in the individual home language groups: 66.2–68.5 (individual range: 57.6–90.0)

APPENDIX B

METHOD AND PROCEDURES, CARD SORT TASKS

Primary school age children

Materials. A set of 24 cards was used, the 2s, 3s, 4s, 8s, 9s, and 10s from all four suits of a normal card deck. The 2s, 3s, and 4s were considered "low" numbers, the 8s, 9s, and 10s the "high" numbers. The four sorts that this group were asked to conduct involved (A) low red cards, (B) high clubs, (C) low diamonds, and (D) low black cards. A stop watch was used for timing response time.

Procedure. The experimenter instructed the child to sort the cards four times, into low red cards, high clubs, low diamonds, and low black cards, in that order or its reverse. Approximately half the children received the sorts in ABCD order, approximately half in DCBA order.

Teens, younger adults, and older adults

Materials. A full set of 52 cards was used. The four sorts that this group were asked to conduct involved (A) odd red cards, (B) even clubs, (C) odd diamonds, and (D) odd black cards. A stop watch was used for timing.

Procedure. The experimenter instructed the participant to sort the cards four times, into (A) odd red cards, (B) even clubs, (C) odd diamonds, and (D) odd black cards. Approximately half the participants received the sorts in ABCD order, approximately half in DCBA order.

Younger children/preschoolers

Materials. A set of 20 cards was used. These cards depicted two types of shapes, circles, and squares ("balls" and "blocks"), and they were of two sizes, small, and large. Children were asked to first sort the cards according to one of these features (shape or size), and then according to the other. A stop watch was used for timing.

Procedure. The child was asked to sort the cards twice, into balls and blocks or into big and little shapes. Approximately half the children received the ball/block sort first, approximately half the big shape/little shape sort first.

APPENDIX C

PROCEDURE FOR THE METALINGUISTIC TASK

Participants heard sentences read to them orally. They were asked to judge whether a sentence was grammatical, and to correct it if it was ungrammatical, with the following instructions:

These are my friends, Sali and Twmi. Sometimes, Twmi gets mixed up when he talks. Sometimes, he doesn't know how to talk very well. Sometimes, he says things right but they're just silly. Sometimes, he says things that make sense, but he says them wrong. And sometimes, he says things that are silly *and* wrong. Sali helps him when he says things the wrong way. She tells Twmi how he should say it. It's okay to say silly things sometimes though, isn't it? So Sali only tells him to say things right when Twmi says them wrong. If he says something that's silly but sounds right, that's ok.

For example, if Twmi says "I am a banana," Sali says "That's very silly but you said it right."

And if Twmi says "My hair be yellow" then Sali says "That makes sense," but it's not right, you should say "my hair *is* yellow"

And if Twmi says "I can talk," Sali says "yes, that's right."

And if Twmi says "I be a lemon," Sali says "That's silly *and* it's wrong." It's okay to be silly, but you should say "I *am* a lemon."

So, if Twmi says (Practice Sentence 1), what do you think Sali will say? (Child responds Right, Silly or Not Right). How should Sali tell Twmi to say it?

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 19 May 2014
doi: 10.3389/fpsyg.2014.00446

ADHD among young adults born at extremely low birth weight: the role of fluid intelligence in childhood

Ayelet Lahat[1]*, Ryan J. Van Lieshout[2], Saroj Saigal[3], Michael H. Boyle[2] and Louis A. Schmidt[1]

[1] *Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON, Canada*
[2] *Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada*
[3] *Department of Pediatrics, McMaster University, Hamilton, ON, Canada*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Peter J. Anderson, Murdoch Childrens Research Institute, Australia
Filippo Dipasquale, University Hospital Policlinico - Vittorio Emanuele, Italy

****Correspondence:***
Ayelet Lahat, Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
e-mail: lahata@mcmaster.ca

Poor executive function (EF) has been linked to attention-deficit/hyperactivity disorder (ADHD). Children born at extremely low birth weight (ELBW; <1000 g) have been found to show both poor EF, as well as elevated levels of symptoms of ADHD. In the present study, we examined whether fluid intelligence moderates the link between birth weight and later ADHD symptoms by prospectively following a cohort of 179 survivors who were born at ELBW. When participants were 8 years-old, they were matched with 145 normal birth weight (NBW; ≥2500 g) control participants. At age 8, fluid intelligence was measured, and during young adulthood (ages 22–26), participants' self-reported levels of ADHD symptoms were examined. We found that ELBW survivors, who also showed poor fluid intelligence, had the highest rates of ADHD symptoms, and particularly, symptoms of inattention. These findings point to the importance of examining developmental trajectories that contribute to risk for psychopathology in those exposed to intrauterine adversity.

Keywords: executive function, fluid intelligence, ADHD, ELBW, longitudinal

INTRODUCTION

Extremely low birth weight (ELBW; <1000 g) survivors are among the tiniest and most vulnerable babies. Compared to individuals born at normal birth weight (NBW; ≥2500 g), those born at very low birth weight (VLBW; <1500 g) and smaller have been found to be at increased risk for later psychopathology, including Attention Deficit/Hyperactivity Disorder (ADHD; Szatmari et al., 1990, 1993; Ross et al., 1991; Botting et al., 1997; Whitaker et al., 1997; Taylor et al., 2000; Bhutta et al., 2002; Elgen et al., 2002; Foulder-Hughes and Cooke, 2003; Indredavik et al., 2004; Strang-Karlsson et al., 2008; Hack et al., 2009; Johnson et al., 2010; Johnson and Marlow, 2011). However, not all ELBW survivors develop ADHD, and very little is known about the developmental trajectories that lead to risk and resilience among these individuals. Accordingly, the aim of the present study was to examine the role of executive function (EF), and specifically fluid intelligence, which may serve as a putative mechanism underlying variation in ADHD risk among individuals born at ELBW.

It is important to point out that although not all low birth weight babies are born prematurely, most babies born ELBW and VLBW are. Premature birth may be associated with a greater risk for symptoms of inattention than hyperactivity/impulsivity, and some studies have reported higher rates of the inattentive subtype of ADHD compared with hyperactive/impulsive subtype in ELBW and VLBW children (Botting et al., 1997; Indredavik et al., 2004; Hack et al., 2009; Johnson et al., 2010). Indeed, some have proposed (Strang-Karlsson et al., 2008) that the ADHD of preterm children is more "pure," as it is characterized by less hyperactivity in relation to inattention, as well as by a more even sex distribution, and it is less frequently accompanied by comorbid disorders (Szatmari et al., 1993; Botting et al., 1997; Elgen et al., 2002; Indredavik et al., 2004). These findings have led some (Szatmari et al., 1990; Hille et al., 2001) to suggest that premature children are susceptible to a more biologically determined form of attention deficit associated with impaired brain growth (Peterson et al., 2000, 2003; Rushe et al., 2001; Kapellou et al., 2006).

For example, Indredavik et al. (2005) found that ADHD symptoms were associated with reduction in white matter volumes and thinning of the corpus callosum in VLBW adolescents. This correlation between symptoms and white matter volume was due primarily to a specific association with inattention scores. In a separate study (Skranes et al., 2007), inattention, but not hyperactivity scores, were associated with fractional anisotropy measurements of white matter in VLBW adolescents. Such white matter abnormalities are associated with difficulties in EF (Edgin et al., 2008), the control over thought and action in situations that require problem solving (Zelazo et al., 2008). Thus, impairments in the underlying cognitive mechanisms that are associated with these structural brain differences, such as EF, could have important implications for later developmental outcomes.

EF also has been referred to as fluid intelligence (Blair, 2006). Although the relation between EF and fluid intelligence has been debated (Birney et al., 2006; Burgess et al., 2006; Garlick and Sejnowski, 2006; Heitz et al., 2006), both entail cognitive processing not necessarily associated with any specific content domain, and involve the active or effortful maintenance of information in working memory for the purposes of planning and performing goal-directed behavior (Kane and Engle, 2002). Since domain general indicators of cognitive abilities involve functions such as information maintenance, attention shifting, and resistance to interference—measures of fluid intelligence have demonstrated significant associations with

performance on measures of general intelligence (Embretson, 1995; Engle et al., 1999). However, there is also evidence for a dissociation between fluid intelligence and general intelligence (see Blair, 2006, for a review), namely, fluid intelligence seems to be a specific subset of more global cognitive abilities (Séguin and Zelazo, 2005).

Impairments in EF/fluid intelligence, particularly in the domains of response inhibition, planning, vigilance, and working memory have been associated with ADHD (see Pennington and Ozonoff, 1996; Willcutt et al., 2005). Studies examining the association between EF and ADHD suggest that poor EF is primarily associated with inattentive symptoms of ADHD rather than hyperactivity or impulsivity (Chhabildas et al., 2001; Nigg et al., 2005; Willcutt et al., 2005). Given that inattentive symptoms of ADHD are more prevalent among individuals born prematurely, it is likely that fluid intelligence plays a role in the development of ADHD among individuals born at ELBW. Indeed, Nadeau et al. (2001) observed that general cognitive ability mediated the relation between extreme preterm birth and hyperactivity, whereas the relation between extreme preterm birth and inattention was mediated specifically by working memory, a specific type of EF.

Children and adolescents who were born preterm have been found to have poorer EF abilities than those born at term (Anderson and Doyle, 2004; Böhm et al., 2007; Luu et al., 2011; Baron et al., 2012). For example, compared to term controls, adolescents born preterm showed deficits in EF abilities, including verbal fluency, inhibition, cognitive flexibility, planning/organization, and working memory, as well as poorer verbal and visuospatial memory (Luu et al., 2011). Böhm et al. (2007) reported that NBW controls surpassed VLBW children on EF, even after controlling for IQ. In another study, at 3 years of age, children born at ELBW performed more poorly than term-born age-mates on working memory and inhibition tasks and had the highest percentage of incomplete performance on a continuous performance test (Baron et al., 2012). Finally, in a different report comparing 8–9 years old ELBW survivors to their NBW peers (Anderson and Doyle, 2004), EF was reduced in the ELBW group.

Given that not all ELBW survivors go on to develop ADHD, and since poor EF is associated with both ELBW and ADHD, we examined here the moderating role of fluid intelligence in understanding the relation between ELBW and symptoms of ADHD. Thus, we conducted a prospective longitudinal study, which included three different time points: (1) birth, (2) middle childhood (age 8), and (3) young adulthood (ages 22–26). This allowed us to examine a developmental trajectory leading to developmental outcomes. We were interested in predicting from birth into young adulthood, which was our endpoint visit.

The cohort of ELBW survivors was followed-up at age 8 and again at 22–26 years of age. A control sample matched on age, sex, and SES was recruited at 8 years of age. During the 8 years visit, participants completed Raven's Colored Progressive Matrices Test (RCPM; Raven, 1983), a measure of fluid intelligence (Blair, 2006) and the Wechsler Intelligence Scale for Children (WISC-R; Wechsler, 1974), measuring general intelligence. As young adults, participants completed the ADHD Rating Scale (Barkley and Murphy, 1998), and the Young Adult Self Report (YASR; Achenbach, 1997). We expected that fluid intelligence would moderate the link between birth weight group and symptoms of ADHD, such that among participants with poor fluid intelligence, ELBW survivors would have the greatest level of ADHD symptoms. Given that ELBW and EF are both linked to the inattentive sub type of ADHD, we expected to find an interaction between birth weight group and fluid intelligence, particularly for inattentive symptoms of ADHD.

METHODS

PARTICIPANTS

This study followed-up a cohort of 397 predominantly Caucasian infants who were born at ELBW (501–1000 g) between 1977 and 1982 to residents of a geographically defined region in central-west Ontario, Canada. Follow-up assessments were conducted when participants were 8- (childhood) and 22–26- (young adulthood) years old. Of the original 397 infants, 179 (45%) survived to hospital discharge from the NICU. There were 13 late deaths, and 166 survived to young adulthood.

During the young adult visit, data were collected on 142 of the 166 (86%) survivors. Reasons for missing data include loss to follow-up ($N = 9$) and refusal ($N = 8$). An additional seven participants had neurosensory impairments (cerebral palsy, blindness, deafness, mental retardation, and microcephaly) and could not complete the assessments. Of these 142, a total of 125 had complete data on the measures collected at the 8-years visit.

The NBW control group was identified and recruited when they and the ELBW cohort were 8 years old. This group comprised a sample of 145 children born at term according to maternal report, between 1977 and 1981. The control sample was selected from class lists provided by local school boards and group-matched with the ELBW cohort on child age, sex, and socioeconomic status (Saigal et al., 1991). Data were collected on 133 of the 145 control participants. Reasons for missing data include loss to follow-up ($N = 5$) and refusal ($N = 7$). All 133 participants had complete data on the measures collected at the 8-years visit.

Data were examined for outliers and participants with more than ± 2 SDs from the mean were removed from all analyses. These outliers were removed as they can affect the mean dramatically and not represent the majority of the group. This resulted in four ELBW and five NBW participants being dropped from the analyses, and thus the final sample included 121 ELBWs and 128 NBWs.

MEASURES

Raven's Colored Progressive Matrices (RCPM; Raven, 1983)

The RCPM was administered when children were 8 years of age. This is a non-verbal measure of fluid intelligence in which the participant is shown colored illustrations with one part missing. The participant is asked to identify and select the missing element that completes the pattern from six possible choices. This measure has been found to be reliable ($\alpha = 0.81-86$) for children at this age (Carlson and Jensen, 1981).

Wechsler Intelligence Scale for Children—Revised (WISC-R; Wechsler, 1974)

Ten subtests of the WISC-R were administered when children were 8 years of age. Digit span and mazes subtests were not included, and the assessment protocol was 3 h long. From these subtests that were administered, verbal and performance IQ scores were calculated. The Full Scale IQ score was derived from the two subscale scores and used in the analysis.

ADHD rating scale (Barkley and Murphy, 1998)

During the young adult visit, ADHD was measured using the ADHD Rating Scale, a self-administered questionnaire comprised of 18 items rated on a four-point scale from 0 (never or rarely) to 3 (very often; $\alpha = 0.85$) (Barkley and Murphy, 1998). The items in this scale map onto the diagnostic criteria for ADHD, and thus three different scores were derived by summing items from this measure: (1) inattention score, (2) hyperactivity/impulsivity score, and (3) total ADHD score. None of the participants met DSM-IV criteria (i.e., six of nine symptoms must be present to show clinical significance) for either the inattentive or hyperactive/impulsive subtypes of ADHD.

Young Adult Self Report (YASR; Achenbach, 1997)

The YASR was completed during the young adult visit. The YASR contains 130 problem items rated as: 0, not true; 1, somewhat or sometimes true; and 2, very true or often true. Based on experts' ratings of the items' consistency with classifications in DSM-IV (A.P.A., 1994), the items were grouped into five DSM-oriented scales (Achenbach et al., 2005): depressive problems ($\alpha = 0.88$), anxiety problems ($\alpha = 0.77$), avoidant personality problems ($\alpha = 0.76$), ADHD problems ($\alpha = 0.72$) and antisocial personality problems ($\alpha = 0.80$); and two higher-order scales: internalizing problems ($\alpha = 0.93$) and externalizing problems ($\alpha = 0.85$). In addition, the YASR can be scored according to syndrome and problems scales and in order to obtain a better understanding of inattention, we report data from the Attention Problems scale. Including this scale, in addition to the ADHD Rating scale, allows examining the same construct with various measures in order to obtain a better understanding of inattention.

RESULTS

DESCRIPTIVE STATISTICS

In order to examine associations between birth weight group and the variables in the study, a series of *t*-tests comparing ELBW and NBW participants were carried out on measures reflecting demographics and SES (sex, mother's highest level of education, and young adult's highest level of education), as well as the main moderator and outcome variables (fluid intelligence, general intelligence, and scores pertaining to ADHD). Mother's highest level of education was measured according to the following rating scale: 1 = No schooling, 2 = Some primary schooling, 3 = Completed primary school, 4 = Some secondary schooling, 5 = Completed secondary school, 6 = Some community college, 7 = Completed community college, 8 = Some university, 9 = Completed university. Young adult's highest level of education was measured according to the following rating scale: 1 = Less than 7th grade, 2 = Junior high school (9th grade), 3 = Partial high school (10 or 11th grade), 4 = High school graduate, 5 = Partial college (at least 1 year or specialized training), 6 = Standard college or university graduation, 7 = Graduate professional training (MSc, MD, MBA, PhD). Descriptive statistics are presented in **Table 1**.

No significant differences were found between the ELBW and NBW groups on sex, mother's highest level of education, and young adult's highest level of education, all *p*s > 0.09. Significant differences were observed between the two groups on birth weight, $[t_{(143.89)} = -56.43, p < 0.0001]$, fluid intelligence, $[t_{(246.67)} = -4.27, p < 0.0001]$, general intelligence, $[t_{(235.76)} = -6.58, p < 0.0001]$, and all WISC-R subtests, all *p*s < 0.005, with NBWs scoring higher on all of these measures. However, ELBW and NBW participants did not significantly differ on ADHD total score, ADHD inattentive score, ADHD hyperactivity/impulsivity score, or YASR attention problems, all *p*s> 0.09. Finally, Pearson correlations revealed that the main variables of interest—fluid intelligence and the various ADHD scores—were not related to one another, all *p*s < 0.59. This result suggests that

Table 1 | Means (and *SD*s) on variables of interest by birth weight group.

	ELBW	NBW
Sex	54 males; 67 females	56 males; 72 females
Birth weight in grams	841.91 (123.91)*	3380.10 (492.63)*
Mother's highest level of education	5.61 (1.91)	6.02 (1.98)
Young adult's highest level of education	4.42 (1.33)	4.70 (1.37)
Full scale WISC-R	92.41 (14.66)*	103.78 (12.44)*
Verbal IQ WISC-R	91.74 (14.23)*	101.30 (12.85)*
Performance IQ WISC-R	94.83 (16.05)*	106.02 (12.90)*
General information WISC-R	9.16 (3.07)*	10.11 (2.80)*
Similarities WISC-R	9.16 (3.38)*	10.66 (3.13)*
Arithmetic WISC-R	8.12 (2.68)*	10.53 (2.66)*
Vocabulary WISC-R	8.45 (3.08)*	10.01 (2.94)*
Comprehension WISC-R	8.63 (2.68)*	10.08 (2.42)*
Picture completion WISC-R	9.12 (2.68)*	11.40 (2.26)*
Picture arrangement WISC-R	9.74 (4.12)*	11.28 (3.00)*
Block design WISC-R	9.21 (3.11)*	11.13 (2.96)*
Object assembly WISC-R	9.41 (3.06)*	10.90 (2.76)*
Coding WISC-R	8.71 (3.38)*	9.83 (2.65)*
RCPM	40.05 (26.53)*	55.11 (29.12)*
ADHD rating scale inattentive score	3.70 (2.97)	3.06 (2.88)
ADHD rating scale hyperactivity/impulsivity score	4.34 (2.97)	4.38 (2.91)
ADHD rating scale total score	8.04 (5.41)	7.45 (5.10)
YASR attention problems	2.83 (2.35)	2.45 (2.05)

*$p < 0.005$.

fluid intelligence at age 8 is not simply an early presentation of later ADHD.

EFFECTS OF BIRTH WEIGHT GROUP AND FLUID INTELLIGENCE ON ADHD

In order to examine the moderating role of fluid intelligence in understanding the relation between birth weight group and ADHD symptoms, four separate hierarchical multiple regression analyses were carried out. Each regression included the following outcome variables: ADHD total score, ADHD inattentive score, ADHD hyperactivity/impulsivity score, and YASR attention problems. To reduce multi-collinearity and aid in interpretation, mean centered predictors were used. Next, the interaction terms were computed as the product between birth weight group and the mean-centered measure of fluid intelligence. Given links between fluid intelligence and general intelligence (Embretson, 1995; Engle et al., 1999), general intelligence was included as a covariate in each regression. Thus, the first step of each regression analysis included the main effects of full scale WISC-R, birth weight group, and RCPM score. To test for the moderating effect of fluid intelligence on the link between birth weight group and ADHD symptoms, the interaction product term between birth weight group and RCPM score was entered in the second step. Although the regression models and the terms contained in them were examined for significance, the moderation hypothesis was tested by examining whether the second step significantly increased the variance explained by each model. Interactions were probed and plotted according to guidelines by Aiken and West (1991), such that high and low levels of fluid intelligence were defined as $\pm 1SD$. Follow-up statistical tests from these probes are reported below.

For ADHD total score, the interaction between birth weight group and fluid intelligence significantly improved the fit of the model, $\Delta R^2 = 0.02$, $[F_{(1,\ 244)} = 4.54, p < 0.05]$ (see **Table 2** and **Figure 1**). To decompose this interaction, follow-up regressions were conducted. The findings indicate that among participants who had low fluid intelligence, birth weight group was related to ADHD score $\beta = 0.20$, $[t_{(244)} = 2.10, p < 0.05]$, such that participants who were at greatest risk at birth (i.e., ELBWs), had the highest ADHD score. However, no such relation emerged in participants with high fluid intelligence, $\beta = -0.09$, $[t_{(244)} = -0.88, p = 0.38]$.

For ADHD inattentive score, the interaction between birth weight group and fluid intelligence significantly improved the fit of the model, $\Delta R^2 = 0.02$, $[F_{(1,244)} = 5.72, p < 0.05]$ (see **Table 2**). To decompose this interaction, follow-up regressions were conducted. The findings indicate that among participants who had low fluid intelligence, birth weight group was related to ADHD inattentive score $\beta = 0.24$, $[t_{(244)} = 2.58, p < 0.01]$, such that participants who were at greatest risk at birth (i.e., ELBWs), had the highest ADHD inattentive score. However, no such relation emerged in participants with high fluid intelligence, $\beta = -0.07$, $[t_{(244)} = -0.76, p = 0.45]$.

When predicting YASR attention problems a trend was found for the interaction between birth weight group and fluid intelligence, $\Delta R^2 = 0.01$, $[F_{(1,\ 244)} = 3.46, p = 0.06]$ (see **Table 2**). To decompose this interaction, follow-up regressions were conducted. The findings indicate that among participants who had low fluid intelligence, birth weight group was related to YASR attention problems $\beta = 0.21$, $[t_{(244)} = 2.24, p < 0.05]$, such that participants who were at greatest risk at birth (i.e., ELBWs), had the most YASR attention problems. However, no such relation emerged in participants with high fluid intelligence, $\beta = -0.04$, $[t_{(244)} = -0.36, p = 0.71]$.

Finally, the regression model was not significant for ADHD hyperactivity/impulsivity score, $\Delta R^2 = 0.01$, $[F_{(1,\ 244)} = 2.06, p = 0.15]$.

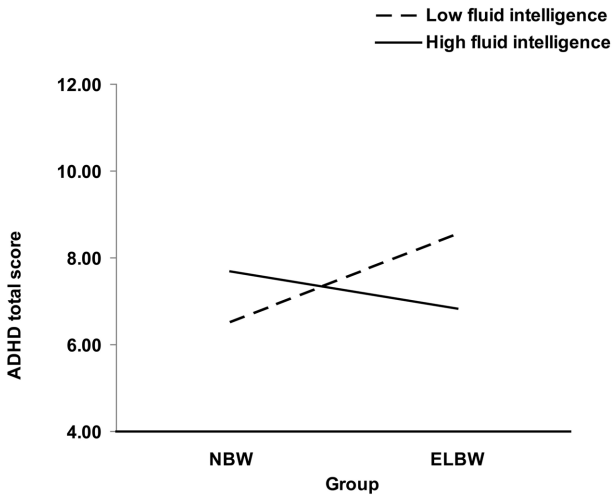


FIGURE 1 | Joint effect of birth weight group and fluid intelligence at age 8 on ADHD total score in young adulthood.

DISCUSSION

This longitudinal study prospectively followed a cohort of ELBW survivors and a matched NBW control sample at age 8 and again during young adulthood (22–26). During the 8-year visit, participants completed measures of fluid and general intelligence. Approximately 15 years later, as young adults, participants provided self-report of ADHD symptoms. As predicted, fluid intelligence moderated the link between birth weight and ADHD symptoms. In particular, ELBW survivors with poor fluid intelligence were at the greatest risk for later ADHD symptoms, particularly symptoms pertaining to the inattentive sub type of ADHD. These findings suggest that fluid intelligence is an important mechanism involved in developmental trajectories that lead ELBW survivors to develop later symptoms of ADHD.

Importantly, our analyses for ADHD total score and ADHD inattentive score were statistically significant even when general intelligence was included in the models as a covariate. This result suggests that fluid intelligence plays a specific role in the association between birth weight group and ADHD above and beyond the role of general intelligence. This finding is important given the presence of substantial correlations between fluid intelligence and measures of general intelligence (e.g., Embretson, 1995; Engle et al., 1999).

Our findings are also consistent with work on VLBW participants, suggesting that links between extreme preterm birth and

Table 2 | Hierarchical multiple regression analysis predicting ADHD symptoms.

Variables by step	ADHD total score		Inattentive score		YASR attention problems	
	R^2	β (t)	R^2	β (t)	R^2	β (t)
Step 1 (df 3/245)	0.004		0.02		−0.004	
WISC-R		−0.004 (−0.04)		−0.09 (−1.01)		0.03 (0.30)
RCPM		0.02 (0.23)		0.05 (0.62)		−0.03 (−0.33)
Birth weight group		0.06 (0.87)		0.09 (1.28)		0.09 (1.33)
Step 2 (df 4/244)	0.02*		0.04*		0.006	
WISC-R		0.004 (0.05)		−0.08 (−0.91)		0.03 (0.38)
RCPM		0.14 (1.36)		0.19 (1.84)		0.08 (0.75)
Birth weight group		0.06 (0.83)		0.08 (1.24)		0.09 (1.29)
Birth weight group × RCPM		−0.18 (−2.13)*		−0.21 (−2.39)*		−0.16 (−1.86)

*$p < 0.05$.

inattention was mediated by working memory, a specific type of EF (Nadeau et al., 2001). Our findings extend this previous research in two important ways: (1) using a moderation approach, we extended this work to ELBW survivors, and (2) our longitudinal study extended over a longer period of time, such that birth weight and fluid intelligence at age 8 interacted to predict ADHD symptoms in young adulthood. Therefore, our findings make a major contribution to research examining developmental trajectories leading to negative outcomes among survivors born at the most severe levels of early adversity.

The present study also extends previous work by Boyle et al. (2011) who found no evidence of group differences in ADHD symptoms during young adulthood with the same cohort of ELBW survivors and controls. This finding was replicated in the present study when directly comparing the two groups. However, the hierarchical regressions revealed that the link between birth weight group and ADHD is more complex, with an interaction between birth weight group and fluid intelligence in predicting later ADHD symptoms, and symptoms of inattention in particular. It should be noted that the interaction between birth weight group and fluid intelligence explained only a small amount of the variance in the ADHD variables examined. This suggests that there are other factors involved in adult ADHD symptoms that were not measured in the present study.

Several authors have proposed that symptoms of ADHD arise from a primary deficit in EF (Pennington and Ozonoff, 1996; Barkley, 1997; Schachar et al., 2000; Castellanos and Tannock, 2002), or that poor EF is an earlier presentation of ADHD. However, in the present study, we did not find a direct relation between fluid intelligence at age 8 and ADHD symptoms during young adulthood. This finding is in line with Willcutt et al.'s (2005) argument that difficulties with EF appear to be only one of many important components of the complex neuropsychology of ADHD.

In the present study, fluid intelligence moderated ADHD symptoms using the total ADHD score on the ADHD Rating Scale (Barkley and Murphy, 1998), as well as the inattentive score on this scale. Furthermore, we observed a trend for Attention Problems using the YASR. These findings are consistent with previous research suggesting associations between birth weight and ADHD (Botting et al., 1997; Indredavik et al., 2004; Hack et al., 2009; Johnson et al., 2010), as well as EF and ADHD (Chhabildas et al., 2001; Nigg et al., 2005; Willcutt et al., 2005), particularly for the inattentive ADHD subtype. Our findings suggest that it is the combination of *both* being born at ELBW *and* having poor fluid intelligence that together contribute to the prediction of later ADHD symptoms, and particularly symptoms of inattention.

It is important to note that the RCPM is only one of many fluid intelligence measures. In addition, although fluid intelligence and EF involve the same underlying processes (Blair, 2006), there is some debate about equating these two constructs (Birney et al., 2006; Burgess et al., 2006; Garlick and Sejnowski, 2006; Heitz et al., 2006). For example, some argue that working memory and fluid intelligence are highly related but separable, and suggest that the mechanism behind the relation is controlled attention—an ability that is dependent on normal functioning of the prefrontal cortex (Heitz et al., 2006).

In summary, the present study followed prospectively the oldest known cohort of ELBW survivors and a matched control sample over a period of 26 years. Fluid intelligence was assessed at 8 years of age, and ADHD symptoms were assessed at 22 to 26 years of age. Our findings indicate that among individuals with poor fluid intelligence measured at age 8, ELBW survivors had the highest level of ADHD symptoms as young adults. These findings point to the importance of examining possible moderating mechanisms that contribute to developmental outcomes and risk for psychopathology.

ACKNOWLEDGMENTS

This research was supported by a Canadian Institutes of Health Research Team Grant (awarded to Louis A. Schmidt), a National Institute of Child Health and Human Development operating grant (awarded to Saroj Saigal), and a Social Sciences and Humanities Research Council of Canada Banting Post Doctoral Fellowship (awarded to Ayelet Lahat). We wish to thank the study participants and their families for their continued participation in this work, as well as Lorraine Hoult and Barbara Stoskopf for their help with data collection and data entry at the visits reported herein, and Keeth Krishnan for his help with data analysis.

REFERENCES

Achenbach, T. M. (1997). *Manual for the Young Adult Self-Report and Young Adult Behavior Checklist.* Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M., Bernstein, A., and Dumenci, L. (2005). DSM oriented scales and statistically based syndromes for ages 18 to 59: Linking taxonomic paradigms to facilitate multitaxonomic approaches. *J. Pers. Assess.* 84, 49–63. doi: 10.1207/s15327752jpa8401_10

Aiken, L. S., and West, G. M. (1991). *Multiple regression: Testing and Interpreting Interactions.* Newbury Park, CA: Sage.

Anderson, P. J., and Doyle, L. W. (2004). Executive functioning in school-aged children who were born very preterm or with extremely low birth weight in the 1990s. *Pediatrics* 114, 50–57. doi: 10.1542/peds.114.1.50

A.P.A. (1994). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* 121, 65–94. doi: 10.1037/0033-2909.121.1.65

Barkley, R. A., and Murphy, K. R. (1998). *Attention-Deficit Hyperactivity Disorder: a Clinical Workbook.* New York, NY: Guilford Press.

Baron, I. S., Kerns, K. A., Muller, U., Ahronovich, M. D., and Litman, F. R. (2012). Executive functions in extremely low birth weight and late-preterm preschoolers: effects on working memory and response inhibition. *Child Neuropsychol.* 18, 586–599. doi: 10.1080/09297049.2011.631906

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., and Anand, K. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *J. Am. Med. Assoc.* 288, 728–737. doi: 10.1001/jama.288.6.728

Birney, D. P., Bowman, D. B., and Pallier, G. (2006). Prior to paradigm integration, the task is to resolve construct definitions of gF and WM. *Behav. Brain Sci.* 29, 127–129. doi: 10.1017/S0140525X06249033

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behav. Brain Sci.* 29, 109–124. doi: 10.1017/S0140525X06009034

Böhm, B., Smedler, A. C., and Forssberg, H. (2007). Impulse control, working memory and other executive functions in preterm children when starting school. *Acta Paediatr.* 93, 1363–1371. doi: 10.1111/j.1651-2227.2004.tb02938.x

Botting, N., Powls, A., Cooke, R. W., and Marlow, N. (1997). Attention deficit hyperactivity disorders and other psychiatric outcomes in very low birthweight children at 12 years. *J. Child Psychol. Psychiatry* 38, 931–941. doi: 10.1111/j.1469-7610.1997.tb01612.x

Boyle, M. H., Miskovic, V., Van Lieshout, R. J., Duncan, L., Schmidt, L. A., Hoult, L., et al. (2011). Psychopathology in young adults born at extremely low birth weight. *Psychol. Med.* 41, 1763–1774. doi: 10.1017/S0033291710002357

Burgess, G. C., Braver, T. S., and Gray, J. R. (2006). Exactly how are fluid intelligence, working memory, and executive function related? Cognitive neuroscience approaches to investigating the mechanisms of fluid cognition. *Behav. Brain Sci.* 29, 128–129. doi: 10.1017/S0140525X0625903X

Carlson, J. S., and Jensen, C. M. (1981). Reliability of the raven colored progressive matrices test: age and ethnic group comparisons. *J. Consult. Clin. Psychol.* 49, 320. doi: 10.1037/0022-006X.49.3.320

Castellanos, F. X., and Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nat. Rev. Neurosci.* 3, 617–628. doi: 10.1038/nrn896

Chhabildas, N., Pennington, B. F., and Willcutt, E. G. (2001). A comparison of the neuropsychological profiles of the DSM-IV subtypes of ADHD. *J. Abnorm. Child Psychol.* 29, 529–540. doi: 10.1023/A:1012281226028

Edgin, J. O., Inder, T. E., Anderson, P. J., Hood, K. M., Clark, C. A., and Woodward, L. J. (2008). Executive functioning in preschool children born very preterm: relationship with early white matter pathology. *J. Int. Neuropsychol. Soc.* 14, 90–101. doi: 10.1017/S1355617708080053

Elgen, I., Sommerfelt, K., and Markestad, T. (2002). Population based, controlled study of behavioural problems and psychiatric disorders in low birthweight children at 11 years of age. *Arch. Dis. Child. Fetal Neonatal. Ed.* 87, F128–F132. doi: 10.1136/fn.87.2.F128

Embretson, S. E. (1995). The role of working memory capacity and general control processes in intelligence. *Intelligence* 20, 169–189. doi: 10.1016/0160-2896(95)90031-4

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128, 309–331. doi: 10.1037/0096-3445.128.3.309

Foulder-Hughes, L. A., and Cooke, R. W. (2003). Motor, cognitive, and behavioural disorders in children born very preterm. *Dev. Med. Child Neurol.* 45, 97–103. doi: 10.1111/j.1469-8749.2003.tb00912.x

Garlick, D., and Sejnowski, T. J. (2006). There is more to fluid intelligence than working memory capacity and executive function. *Behavioral and Brain Sciences* 29, 134–135. doi: 10.1017/S0140525X0630903X

Hack, M., Taylor, H. G., Schluchter, M., Andreias, L., Drotar, D., and Klein, N. (2009). Behavioral outcomes of extremely low birth weight children at age 8 years. *J. Dev. Behav. Pediatr.* 30, 122–130. doi: 10.1097/DBP.0b013e31819e6a16

Heitz, R. P., Redick, T. S., Hambrick, D. Z., Kane, M. J., Conway, A. R. A., and Engle, R. W. (2006). Working memory, executive function, and general fluid intelligence are not the same. *Behav. Brain Sci.* 29, 135–136. doi: 10.1017/S0140525X06319036

Hille, E., Den Ouden, A. L., Saigal, S., Wolke, D., Lambert, M., Whitaker, A., et al. (2001). Behavioural problems in children who weigh 1000 g or less at birth in four countries. *Lancet* 357, 1641–1643. doi: 10.1016/S0140-6736(00)04818-2

Indredavik, M. S., Skranes, J. S., Vik, T., Heyerdahl, S., Romundstad, P., Myhr, G. E., et al. (2005). Low-birth-weight adolescents: psychiatric symptoms and cerebral MRI abnormalities. *Pediatr. Neurol.* 33, 259–266. doi: 10.1016/j.pediatrneurol.2005.05.002

Indredavik, M. S., Vik, T., Heyerdahl, S., Kulseng, S., Fayers, P., and Brubakk, A. (2004). Psychiatric symptoms and disorders in adolescents with low birth weight. *Arch. Dis. Child. Fetal Neonatal. Ed.* 89, F445–F450. doi: 10.1136/adc.2003.038943

Johnson, S., Hollis, C., Kochhar, P., Hennessy, E., Wolke, D., and Marlow, N. (2010). Psychiatric disorders in extremely preterm children: longitudinal finding at age 11 years in the EPICure study. *J. Am. Acad. Child Adolesc. Psychiatry* 49, 453–463. doi: 10.1016/j.jaac.2010.02.002

Johnson, S., and Marlow, N. (2011). Preterm birth and childhood psychiatric disorders. *Pediatr. Res.* 69, 11R–18R. doi: 10.1203/PDR.0b013e318212faa0

Kane, M. J., and Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychon. Bull. Rev.* 9, 637–671. doi: 10.3758/BF03196323

Kapellou, O., Counsell, S. J., Kennea, N., Dyet, L., Saeed, N., Stark, J., et al. (2006). Abnormal cortical development after premature birth shown by altered allometric scaling of brain growth. *PLoS Med.* 3:e265. doi: 10.1371/journal.pmed.0030265

Luu, T. M., Ment, L., Allan, W., Schneider, K., and Vohr, B. R. (2011). Executive and memory function in adolescents born very preterm. *Pediatrics* 127, e639–e646. doi: 10.1542/peds.2010-1421

Nadeau, L., Boivin, M., Tessier, R., Lefebvre, F., and Robaey, P. (2001). Mediators of behavioral problems in 7-year-old children born after 24 to 28 weeks of gestation. *J. Dev. Behav. Pediatr.* 22, 1–10. doi: 10.1097/00004703-200102000-00001

Nigg, J. T., Stavro, G., Ettenhofer, M., Hambrick, D. Z., Miller, T., and Henderson, J. M. (2005). Executive functions and ADHD in adults: Evidence for selective effects on ADHD symptom domains. *J. Abnorm. Psychol.* 114, 706–717. doi: 10.1037/0021-843X.114.3.706

Pennington, B. F., and Ozonoff, S. (1996). Executive functions and developmental psychopathology. *J. Child Psychol. Psychiatry* 37, 51–87. doi: 10.1111/j.1469-7610.1996.tb01380.x

Peterson, B. S., Anderson, A. W., Ehrenkranz, R., Staib, L. H., Tageldin, M., Colson, E., et al. (2003). Regional brain volumes and their later neurodevelopmental correlates in term and preterm infants. *Pediatrics* 111, 939–948. doi: 10.1542/peds.111.5.939

Peterson, B. S., Vohr, B., Staib, L. H., Cannistraci, C. J., Dolberg, A., Schneider, K. C., et al. (2000). Regional brain volume abnormalities and long-term cognitive outcome in preterm infants. *J. Am. Med. Assoc.* 284, 1939–1947. doi: 10.1001/jama.284.15.1939

Raven, J. (1983). *Guide to Using the Coloured Progressive Matrices.* London: HK Lewis.

Ross, G., Lipper, E. G., and Auld, P. A. (1991). Educational status and school-related abilities of very low birth weight premature children. *Pediatrics* 88, 1125–1134.

Rushe, T., Rifkin, L., Stewart, A., Townsend, J., Roth, S., Wyatt, J., et al. (2001). Neuropsychological outcome at adolescence of very preterm birth

and its relation to brain structure. *Dev. Med. Child Neurol.* 43, 226–233. doi: 10.1017/S0012162201000433

Saigal, S., Szatmari, P., Rosenbaum, P., Campbell, D., and King, S. (1991). Cognitive abilities and school performance of extremely low birth weight children and matched term control children at age 8 years: a regional study. *J. Pediatr.* 118, 751–760. doi: 10.1016/S0022-3476(05)80043-5

Schachar, R., Mota, V. L., Logan, G. D., Tannock, R., and Klim, P. (2000). Confirmation of an inhibitory control deficit in attention-deficit/hyperactivity disorder. *J. Abnorm. Child Psychol.* 28, 227–235. doi: 10.1023/A:1005140103162

Séguin, J. R., and Zelazo, P. D. (2005). "Executive function in early physical aggression," in *Developmental Origins of Aggression,* eds R. E. Tremblay, W. W. Hartup, and J. Archer (New York, NY: Guilford Press), 307–329.

Skranes, J., Vangberg, T., Kulseng, S., Indredavik, M., Evensen, K., Martinussen, M., et al. (2007). Clinical findings and white matter abnormalities seen on diffusion tensor imaging in adolescents with very low birth weight. *Brain* 130, 654–666. doi: 10.1093/brain/awm001

Strang-Karlsson, S., Räikkönen, K., Pesonen, A.-K., Kajantie, E., Paavonen, E., Lahti, J., et al. (2008). Very low birth weight and behavioral symptoms of attention deficit hyperactivity disorder in young adulthood: the Helsinki study of very-low-birth-weight adults. *Am. J. Psychiatry* 165, 1345–1353. doi: 10.1176/appi.ajp.2008.08010085

Szatmari, P., Saigal, S., Rosenbaum, P., and Campbell, D. (1993). Psychopathology and adaptive functioning among extremely low birthweight children at eight years of age. *Dev. Psychopathol.* 5, 345–357. doi: 10.1017/S0954579400004454

Szatmari, P., Saigal, S., Rosenbaum, P., Campbell, D., and King, S. (1990). Psychiatric disorders at five years among children with birthweights < 1000 g: a regional perspective. *Dev. Med. Child Neurol.* 32, 954–962. doi: 10.1111/j.1469-8749.1990.tb08117.x

Taylor, H. G., Klein, N., Minich, N., and Hack, M. (2000). Middle school-age outcomes in children with very low birth-weight. *Child Dev.* 71, 1495–1511. doi: 10.1111/1467-8624.00242

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children - Revised.* New York, NY: Psychological Corp.

Whitaker, A. H., Van Rossem, R., Feldman, J. F., Schonfeld, I. S., Pinto-Martin, J. A., Torre, C., et al. (1997). Psychiatric outcomes in low-birth-weight children at age 6 years: relation to neonatal cranial ultrasound abnormalities. *Arch. Gen. Psychiatry* 54, 847–856. doi: 10.1001/archpsyc.1997.01830210091012

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., and Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biol. Psychiatry* 57, 1336–1346. doi: 10.1016/j.biopsych.2005.02.006

Zelazo, P. D., Carlson, S. M., and Kesek, A. (2008). "Development of executive function in childhood," in *Handbook of Developmental Cognitive Neuroscience, 2nd Edn.,* eds C. A. Nelson and M. Luciana (Cambridge, MA: MIT Press), 553–574.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 February 2014; accepted: 27 April 2014; published online: 19 May 2014.
*Citation: Lahat A, Van Lieshout RJ, Saigal S, Boyle MH and Schmidt LA (2014) ADHD among young adults born at extremely low birth weight: the role of fluid intelligence in childhood. Front. Psychol. **5**:446. doi: 10.3389/fpsyg.2014.00446*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Lahat, Van Lieshout, Saigal, Boyle and Schmidt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 10 June 2014
doi: 10.3389/fpsyg.2014.00566

The development of cognitive control in children with chromosome 22q11.2 deletion syndrome

Heather M. Shapiro[1]*, Flora Tassone[1,2], Nimrah S. Choudhary[2] and Tony J. Simon[1]

[1] Department of Psychiatry and Behavioral Sciences, MIND Institute, University of California at Davis, Sacramento, CA, USA
[2] Department of Biochemistry and Molecular Medicine, University of California at Davis, Sacramento, CA, USA

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Carmelo Mario Vicario, Bangor, UK
Koichi Haishi, Saitama University, Japan

***Correspondence:**
Heather M. Shapiro, MIND Institute, 2825 50th Street, Room 1357, Sacramento, CA 95817, USA
e-mail: hmshapiro@ucdavis.edu

Chromosome 22q11.2 Deletion Syndrome (22q11.2DS) is caused by the most common human microdeletion, and it is associated with cognitive impairments across many domains. While impairments in cognitive control have been described in children with 22q11.2DS, the nature and development of these impairments are not clear. Children with 22q11.2DS and typically developing children (TD) were tested on four well-validated tasks aimed at measuring specific foundational components of cognitive control: response inhibition, cognitive flexibility, and working memory. Molecular assays were also conducted in order to examine genotype of catechol-O-methyltransferase (COMT), a gene located within the deleted region in 22q11.2DS and hypothesized to play a role in cognitive control. Mixed model regression analyses were used to examine group differences, as well as age-related effects on cognitive control component processes in a cross-sectional analysis. Regression models with COMT genotype were also conducted in order to examine potential effects of the different variants of the gene. Response inhibition, cognitive flexibility, and working memory were impaired in children with 22q11.2DS relative to TD children, even after accounting for global intellectual functioning (as measured by full-scale IQ). When compared with TD individuals, children with 22q11.2DS demonstrated atypical age-related patterns of response inhibition and cognitive flexibility. Both groups demonstrated typical age-related associations with working memory. The results of this cross-sectional analysis suggest a specific aberration in the development of systems mediating response inhibition in a sub-set of children with 22q11.2DS. It will be important to follow up with longitudinal analyses to directly examine these developmental trajectories, and correlate neurocognitive variables with clinical and adaptive outcome measures.

Keywords: 22q11.2 deletion syndrome, cognitive control, executive function, childhood cognitive development, developmental disorders, catechol-O-methyltransferase (COMT)

INTRODUCTION

Chromosome 22q11.2 Deletion Syndrome (22q11.2DS) results from a 1.5- to 3-megabase microdeletion on the long (q) arm of chromosome 22 (Carlson et al., 1997) and occurs in approximately one in 2000–4000 live births (Oskarsdóttir et al., 2004; Shprintzen, 2008). Children with this disorder have mild to moderate intellectual impairments (median full scale IQ 70 ± 15) (Scambler, 2000) and a cognitive profile with difficulties on a range of functions including attention and quantitative processing (Simon et al., 2005; Simon, 2008; Simon and Luck, 2011), as well as cognitive control (Bish et al., 2005; Sobin et al., 2005). Importantly, children with 22q11.2DS also have behavioral impairments and are at significantly increased risk for developing schizophrenia in adulthood (Murphy et al., 1999). Approximately 25% of individuals with 22q11.2DS will develop schizophrenia by adulthood (Bassett et al., 2003), rendering it the highest genetic risk factor for the disorder after having a monozygotic twin or two parents with schizophrenia.

In the schizophrenia literature, impairments in cognitive control have been shown to precede symptom onset (Cannon et al., 2003; Brewer et al., 2005; Lencz et al., 2006). There is also evidence for attenuated cognitive control impairments among first-degree relatives of individuals with schizophrenia, suggesting that these deficits might be part of an endophenotype related to genetic susceptibility for the disorder (Snitz et al., 2006). Based on this line of evidence, a better understanding of cognitive control component processes in children with 22q11.2DS, a group with a genetically conferred risk for schizophrenia, might help to identify specific cognitive functions that could act both as biomarkers for conversion risk, and as specific targets for intervention that might reduce that risk.

In the current study, our goal was to take a first step toward characterizing the nature and extent of cognitive control impairments throughout development in children with 22q11.2DS by conducting a cross-sectional analysis in individuals aged 7–14 years. Cognitive control, a term largely synonymous with executive function, describes the dynamic system of mental processes that directs and regulates cognitive resources in order to maximally achieve one's goals. Miyake et al. (2000) described a theoretical framework suggesting that this system encompasses three

foundational cognitive control components, namely response inhibition, cognitive flexibility, and working memory, and that these components are both distinct and interrelated. This system is not static developmentally, but rather each component process has a unique developmental trajectory, and the degree to which the components are distinct or interrelated changes as a function of age (Best and Miller, 2010).

Preliminary evidence suggests that children with 22q11.2DS exhibit impairments in cognitive control processes, as well as neuroanatomical and neurofunctional aberrations in networks believed to support cognitive control processes. In a study aimed at understanding schizophrenia-like cognitive impairments in children with 22q11.2DS aged 7–16 years, Lewandowski et al. (2007) found that performance on a Wisconsin Card Sort task, a well-established paradigm for examining cognitive flexibility, was impaired relative to TD, even after controlling for general intellectual function by including IQ as a regressor in the statistical models. By contrast, working memory impairments, as measured by the Children's California Verbal Learning Test (CVLT-C), were not significant after accounting for IQ in the regression models.

Campbell et al. (2010) also tested cognitive control abilities in children with 22q11.2DS, aged 6–16 years. They found that children with 22q11.2DS had significantly impaired cognitive flexibility relative to TD, as measured by the Switch task from the Maudsley Attention and Response Suppression battery, as well as impaired working memory, as measured by the Children's Memory Scale and a Spatial Working Memory task from the Cambridge Neuropsychological Testing Automated Battery. By contrast, they found no between-group differences on a Go/No-Go task, a well-established paradigm for examining response inhibition. Other studies, however, demonstrated that children with 22q11.2DS had inhibitory control impairments on tasks requiring interference control (Bish et al., 2005) and oculomotor inhibition (Sobin et al., 2005).

Thus, it is evident that while cognitive control systems appear to be impaired in 22q11.2DS, the specific nature of these impairments is unclear. There are a number of factors that could account for differences in the previous literature. First of all, some of the cognitive control measures were extracted from psychometrically well-characterized, standardized behavioral testing instruments. While these tests are valuable, they are not as good at isolating specific cognitive processes, as are experimental neurocognitive tests. Additionally, the previous studies characterized large age ranges, throughout which cognitive control processes are dynamically changing as a function of brain developmental processes. Thus, given the relevance of these impairments to cognitive function in 22q11.2DS, as well as to schizophrenia risk, it is important to characterize the nature and developmental trajectory of cognitive control processes using most sensitive, specific neurocognitive tests of cognitive control component processes.

Preliminary evidence from a cross-sectional sample of individuals aged 7–14 years reported that children with 22q11.2DS had an age-related impairment in the executive control of attention, specifically with respect to a flanker inhibition paradigm (Stoddard et al., 2011). Interestingly, another cross-sectional study examining a complementary aspect of attention, namely attentional orienting, in the same age range demonstrated the opposite pattern: performance in older individuals with 22q11.2DS was significantly better and less variable than that of their younger counterparts (Shapiro et al., 2012). This pattern suggests that different systems of attention and their underlying neural networks are developing with different trajectories in 22q11.2DS. Importantly, it appears that impairments in cognitive control, not general cognitive or attentional function, are preceding the risk period, and might contribute to part of a risk profile. Testing this hypothesis is important for understanding networks that might be particularly plastic in a critical age period during which aberrant neurodevelopment might render a subset of individuals at increased risk for developing schizophrenia.

Here we tested an age range of children with 22q11.2DS and TD comparison children aged 7–14 years on a battery of specific cognitive control component processes for a cross-sectional analysis of the development of cognitive control in this population. Based on Miyake et al.'s (2000) theoretical model of cognitive control foundational components, we examined response inhibition, cognitive flexibility, and working memory using a battery of child-adapted, well-validated neurocognitive tasks to probe each component. Response inhibition was assessed with a canonical stroop task (Stroop, 1935). The second task for measuring response inhibition was a child-friendly "whack-a-mole" version of a Go/No-Go task. Go/No-Go tasks have been widely used in both typically and atypically developing children to examine inhibitory control (Casey et al., 1997). Here participants responded to a frequently occurring target ("Go" trial), and inhibited the pre-potent response to an infrequent target ("No-Go" trial). Cognitive flexibility was examined using a Visually-Cued Card Sort (VCCS), a downward extension of the Wisconsin Card Sort that is geared toward children (Zelazo et al., 2004). In this study participants sorted cards according to rules about shape or color, and the sorting rules changed according to certain criteria. In contrast to the Wisconsin Card Sort, participants received an explicit visual cue indicating the specific rule set by which to sort. Finally a Self-Ordered Pointing Test (SOPT) was used to examine working memory (Petrides and Milner, 1982). Participants identified and responded to a sequence of images, remembering which images they have previously chosen, and select a new image on each subsequent trial.

Beyond age-related associations with cognitive control, we wanted to examine additional factors that might contribute to cognitive control performance in children with 22q11.2DS. The gene for catechol-O-methyltransferase (COMT) is located within the deleted region in 22q11.2DS and is an important regulator of prefrontal dopamine (DA), a neurotransmitter that has previously been reported to play a role in higher-level cognitive processes (Kimberg and D'Esposito, 2003). Given that children with 22q11.2DS have only a single copy of the COMT allele, it is likely that DA modulation is abnormal in these individuals. Importantly, the COMT gene contains two different allelic variations: Val and Met for high and low enzymatic activity, respectively. Previous studies of COMT genotype in 22q11.2DS have yielded differential results, with some studies reporting Met hemizygosity of COMT to be related to poorer outcome on tasks requiring executive control (Baker et al., 2005; Takarae et al., 2009), and others reporting better outcomes (Bearden et al., 2004;

Shashi et al., 2006). Additional studies have found no relationship between COMT genotype and measures of cognitive control in 22q11.2DS (Glaser et al., 2006; Campbell et al., 2010). Thus, in order to investigate this relationship further, we examined cognitive control performance of the participants in this study as a function of COMT variant.

Based on previous evidence of cognitive control impairments in 22q11.2DS, we hypothesized that individuals with the disorder would perform more poorly on the cognitive control tasks relative to TD comparison children. Additionally, we hypothesized that a cross-sectional analysis of cognitive control development would reveal atypical developmental trajectories of specific cognitive control components, with worse performance in older but not younger children with 22q11.2DS, and that this pattern would be true in some but not all of the children.

MATERIALS AND METHODS

PARTICIPANTS

Seventy-one children with chromosome 22q11.2 deletion syndrome (mean age = 11.4[2.5] years; 31 female and 40 male) and 52 typically developing (TD) comparison children (mean age = 10.6[2.2] years; 27 female and 25 male), from 7 to 14 years of age, participated in the study. Data on IQ from the Wechsler Intelligence Scale for Children—4th edition (WISC-IV) (Wechsler, 2003) or the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999) was available from a subset of participants: 55 children with 22q11.2DS and 38 TD participants. Full-scale IQ (FSIQ) ranged from 46 to 103 for children with 22q11.2DS and 80 to 154 for TD children. Biological samples were available for genotyping on 58 of the children with 22q11.2DS. Of these individuals, 31 were hemizygous for the COMT Val allele and 27 were hemizygous for the COMT Met allele. A subsample of the study participants (12 with 22q11.2DS and 8 TD) performed the cognitive task battery at a conference where they did not complete the WASI or submit biological samples, thus contributing to incomplete IQ and COMT data, respectively. Exclusion criteria for both groups included head injury or other focal neurological abnormality. Exclusion criteria for TD participants were the presence of any other learning or behavioral/psychiatric disorder. Additional exclusion criteria on an individual task basis are described under the description for each task below. One participant with 22q11.2DS met exclusion criteria for all tasks and was removed from analysis, resulting in the final sample of 71 children with 22q11.2DS and 52 TD children that are described here. The parents of all participants provided written informed consent based on protocols approved by the Institutional Review Board at the University of California, Davis. **Table 1** depicts the demographic information for children in each group.

Table 1 | Demographic data on children with 22q11.2DS (22q) and TD children.

Cognitive task		*N*	Age in years: Mean (*SD*)	FSIQ: Mean (*SD*)	Gender: Sample size (Male/ Female)	COMT: Sample size (Val/ Met)
Stroop	22q	39	11.4 (2.5)	74.6 (14.2)	21/18	21/14
	TD	29	10.5 (2.1)	114.0 (12.1)	17/12	NA
Go/No-Go	22q	64	11.5 (2.6)	73.9 (12.7)	37/27	28/25
	TD	49	10.7 (2.2)	113.2 (13.7)	24/25	NA
VCCS	22q	62	11.6 (2.5)	73.7 (13.2)	34/28	28/21
	TD	50	10.7 (2.2)	113.2 (13.8)	24/26	NA
SOPT (Verbal)	22q	65	11.5 (2.6)	73.8 (13.4)	36/28	28/24
	TD	52	10.6 (2.2)	112.7 (13.6)	25/27	NA
SOPT (Non-verbal)	22q	56	11.6 (2.6)	73.0 (11.9)	32/24	25/19
	TD	47	10.8 (2.2)	112.5 (13.4)	23/24	NA

MOLECULAR ANALYSES

Genomic DNA was isolated from 3 ml of peripheral blood leukocytes using standard procedure (Qiagen, Valencia, CA). Genotyping analysis for the COMT $Val^{108/158}$ Met was carried out by TaqMan SNP Genotyping Assay (rs4680; Applied Biosystems, Foster City, CA). PCR reaction contained COMT SNP genotyping assay mix, TaqMan master mix and 25 ng DNA per reaction. PCR conditions were 95°C for 10 min, followed by 40 Cycles of 92°C for 15 s and 60°C for 1 min. Allelic discrimination plate read was performed on an Applied Biosystems Real-Time PCR System using the Sequence Detection System (SDS) Software.

TASK PROCEDURE

All participants completed paradigms testing cognitive control component processes, including response inhibition, a cognitive flexibility, and working memory. Tasks were administered on the same Elo 1715L Desktop Touch monitor for all participants.

Response inhibition paradigms

To examine response inhibition, participants completed a computerized version of the canonical Stroop task (Stroop, 1935). Participants were presented with stimuli on a monitor and asked to respond (by pressing one of three colored buttons) in which color font the stimulus was presented (red, green, or blue). In the congruent condition, participants were presented with the words "red," "green," or "blue" in the same font color as the presented word. In the incongruent condition, participants were presented with one of the same three color words; however, the word was presented in a font color that was *different* from the specified color word (**Figure 1A**). There were a total of 240 trials, with 168 and 72 congruent and incongruent trials, respectively. The rationale for this 70–30 congruent-incongruent ratio was to maintain the potency of the rule set for responding to the congruent color. Stimuli were presented for 2000 ms, or until the participant responded, with interstimulus intervals of 200, 500, or 750 ms. The dependent variable here was median response time (RT) on congruent relative to incongruent trials that were preceded by congruent vs. incongruent trials, respectively. Participants were excluded if they performed worse than chance (66.6% accuracy) on congruent or incongruent trials. Seven children with 22q11.2DS were excluded on this basis. This task was completed on a slightly smaller sample of participants (39 participants with

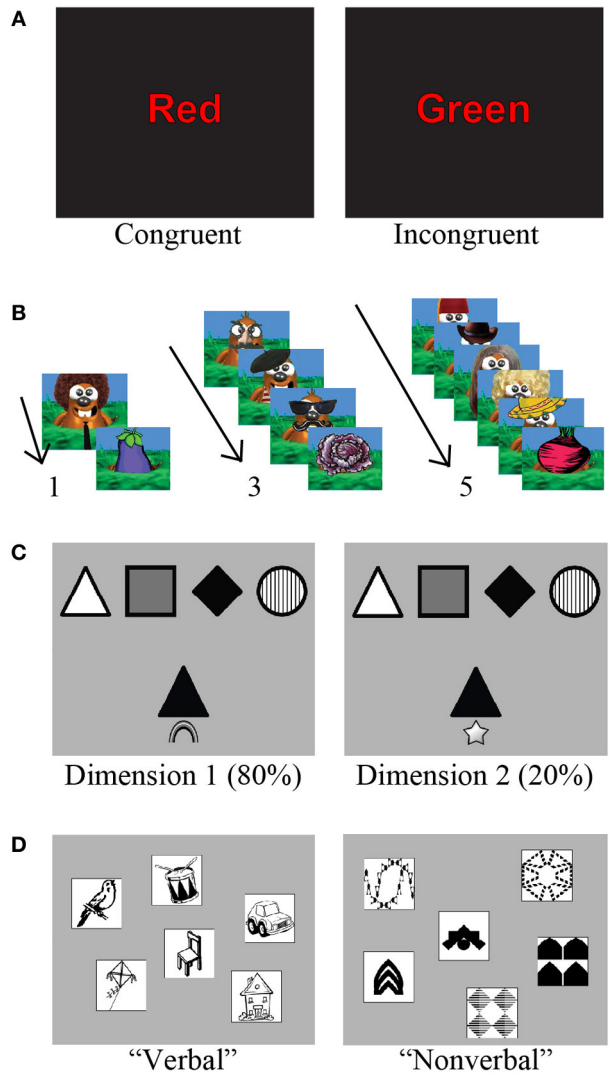


FIGURE 1 | Examples of the cognitive control battery. (A) For the Stroop task, participants had to respond by indicating the ink color on congruent (left) or incongruent (right) trials. **(B)** "Whack-a-mole" Go/No-Go task. Children were instructed to press a button as quickly as possible when a cartoon mole appeared (Go trial), but to avoid pressing the button when a vegetable appeared (No-Go trial). No-Go trials were preceded by 1, 3, or 5 Go trials (Adapted from Shapiro et al., 2013). **(C)** The two gray squares each represent a touch-screen display in the Visually-Cued Card Sort (VCCS). A sample of the target cards can be seen at the top of the screen, while the test card is below. The visual cue appears just below the test card, with a rainbow indicating to sort by color (left panel), and a star indicating to sort by shape (right panel). **(D)** The two gray squares each represent a touch-screen display in the for the verbal and non-verbal Self-Ordered Pointing Test (SOPT), respectively. Images represent a trial with 6 objects, the most difficult condition of the task.

22q11.2DS and 29 TD), due to a modification of the task design that occurred approximately 6 months into the study.

Response inhibition was also measured using a child-adapted version of a Go/No-Go response inhibition task (**Figure 1B**). A subset of this data has been published previously (Shapiro et al., 2013), but our goal here was to extend those findings by including a larger sample of participants, and also examine within-subject differences on this component of the battery relative to the other cognitive control processes. For a full description of the task, please reference Shapiro et al. (2013). Key details are the task parameters including Go (75%) and No-Go (25%) trials. Stimuli were presented for 1000 ms, with interstimulus intervals of 200, 500, or 750 ms. Participants completed 20 trials of each No-Go type (preceded by one, three, or five Go trials, respectively), divided equally into four blocks. Primary outcome measures were accuracy and RT to Go and No-Go trials, respectively. Participants were excluded if they performed at lower than 75% accuracy when responding to the frequently occurring Go stimuli, or outside of 2 standard deviations from the mean for accuracy on No-Go trials. Seven children with 22q11.2DS and three TD participants were excluded on this basis.

Cognitive flexibility paradigm

To examine cognitive flexibility, participants completed a computerized version of the VCCS on a computer with a touch-screen monitor. This is a children's modified version of the Wisconsin card sorting task, and was adapted from a task by Zelazo et al. (2004) that proved to be effective at measuring perseverative behavior in a wide age range of children. At a distance of approximately 60 cm from the computer, participants viewed four target cards that displayed four different shapes (circle, square, diamond, triangle) in four different colors (black, white, gray, striped; **Figure 1C**). They were instructed to sort 50 test cards onto the appropriate target card. The test cards were presented one at a time at a central location beneath the row of target cards. The participants were instructed to sort their cards either by color or by shape, as indicated by the visual cue that appeared below their card. A rainbow was the visual cue that indicated to sort by color, while a star indicated to sort by shape. Forty out of 50 trials were cued to sort by one of the dimensions (color or shape), while the remaining 10 trials were cued to sort by the secondary dimension. For the first 45 participants (23 with 22q11.2DS and 22 TD), color was the primary dimension (Dimension 1), while the remaining 67 participants (39 with 22q11.2DS and 28 TD) were presented with shape as the primary dimension. The trials were uniformly randomized, such that one trial of the secondary sorting dimension (Dimension 2) appeared within every five trials. The participants completed a demonstration of the task, followed by four practice trials, after which they began the 50 test trials. Test cards were presented on the screen for as long as the participants needed to make a response. If the response was incorrect, the test and target cards remained on the screen until the participants selected the correct target, after which the screen refreshed and a new test and target cards were presented. Primary outcome measures were percent accuracy of correctly sorted cards for each dimension (Dimensions 1 and 2 for 80 and 20% frequency, respectively), as well as the ratio of accuracy from Dimension 2 divided by Dimension 1. The ratio score was intended isolate the costing of switching dimensions (i.e., cognitive flexibility) from general card sorting ability on the task. Participants were excluded if their overall performance accuracy was less than 50%, if they did not appear to understand the task after repeating the instructions, or if they did not comply with the task instructions. Nine

children with 22q11.2DS and two TD participants were excluded based on these criteria.

Working memory paradigm

Participants completed a modified version of the SOPT, originally designed by Petrides and Milner (1982). There were two versions of the task: verbal and non-verbal. The verbal version consisted of single-syllable, concretely nameable objects while the non-verbal version involved visual stimuli that were difficult to name or encode verbally. Visual stimuli were chosen from the Dover Clip Art Series, a library of images that is available copyright-free at doverpublications.com. A computer screen displayed an array of images presented on a touch-screen monitor. There were three levels to this task. From easiest to most difficult, the levels involved three, four, or six images, respectively. The most difficult level (six images) can be seen in **Figure 1D**. The participants were asked to point to an object (touch the object on the touch-screen monitor), with the condition that on each subsequent trial they must point to a different object. Each time the participants pointed to an object, the screen refreshed and the relative positions of the images were rearranged at random. Each block consisted of the same number of trials as different objects on the screen. There were four blocks at each level. Primary outcome measures were span (number of correct responses prior to the first error) and number of errors. Participants were excluded if their overall performance accuracy was at chance, if they did not appear to understand the task after repeating the instructions, or if they did not comply with the task instructions. Based on these criteria, six children with 22q11.2DS were excluded from analysis of the verbal version of the task. Fifteen children with 22q11.2DS and five TD children were excluded from analysis of the non-verbal version.

DATA ANALYSIS

Data were processed using scripts written by HS in MatLab (version 7.8) to generate outcome variables from raw data. Mixed model regression analyses were used to determine the effects of between-subject variables (diagnosis group, gender, and testing location) and task variables on primary outcome measures. Age was included as a regressor to examine developmental effects in a cross-sectional analysis. Additional models included full-scale IQ as a regressor in order to assess the relationship of general intellectual abilities with cognitive control function. Finally, COMT genotype was included as a regressor in order to examine the potential relationship of specific genetic variants to the cognitive control processes.

RESULTS

RESPONSE INHIBITION—STROOP TASK

Response inhibition was measured by accuracy and RT on two different Trial Types: congruent and incongruent. Children with 22q11.2DS had overall worse accuracy than TD children [$F_{(1, 65)} = 12.12$, $p = 0.0009$], and there was a trend toward a significant Group × Trial Type interaction, such that children with 22q11.2DS had a relatively worse accuracy on the incongruent relative to congruent trials [$F_{(1, 66)} = 3.30$, $p = 0.07$]. There was no overall group difference in RT [$F_{(1, 65)} = 2.32$, $p = 0.13$], nor a Group × Trial Type interaction in RT [$F_{(1, 66)} = 0.44$, $p = 0.51$].

In order to examine interference effects of preceding trial type, we next examined group performance as a function of four different Trial types: congruent and incongruent trials that were preceded by congruent or incongruent trials, respectively. Thus, the four different Trial Types included: congruent preceded by congruent (*cC*), congruent preceded by incongruent (*iC*), incongruent preceded by congruent (*cI*), and incongruent preceded by incongruent (*iI*). Within each Trial Type, children with 22q11.2DS had significantly worse accuracy relative to TD children across all Trial Types (Supplementary Table 1 and **Figure 2A**). By contrast, there were no group differences in RT for any of the specific trial types (Supplementary Table 1 and **Figure 2B**).

Next, we took the difference of RT on congruent trials that were preceded by incongruent trials (iC), minus that of congruent trials preceded by congruent trials (cC). The goal here was measure the specific interference effects of a prior incongruent trial on congruent RT, relative to RT on a congruent trial that is not preceded by an interfering stimulus (cC). Here we found that children with 22q11.2DS had a significantly larger RT difference (iC – cC) relative to that of TD children [$F_{(1, 65)} = 5.06$, $p = 0.03$], suggesting that this population is more greatly affected by the prior interfering stimulus (Supplementary Table 1 and **Figure 2C**).

RESPONSE INHIBITION—GO/NO-GO TASK

Performance on this task in children with 22q11.2DS has been previously reported (Shapiro et al., 2013) for a subgroup of children with 22q11.2DS ($n = 47$) and of TD children ($n = 36$). Here we report on the results of additional 17 children with 22q11.2DS and 13 TD children. These results are important to report here in

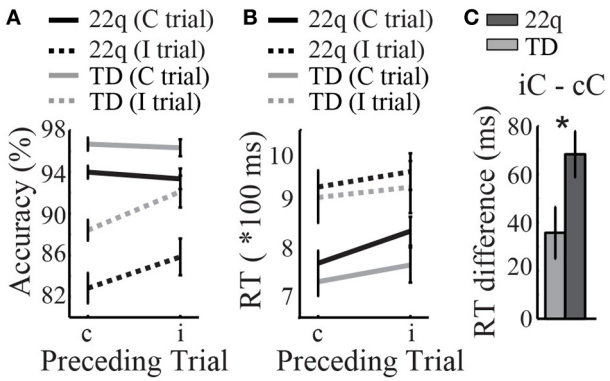


FIGURE 2 | Results of the response inhibition Stroop task. (A) Children with 22q11.2DS had lower accuracy relative to TD participants across all trial types: Congruent (*C*) or Incongruent (*I*) trials preceded by congruent (*c*) or incongruent (*i*) trials, respectively. **(B)** Response time (RT) was similar between groups. **(C)** Children with 22q11.2DS were more greatly affected by a preceding interfering stimulus, as measured by a larger RT difference on congruent trials preceded by incongruent trials (*iC*) relative to congruent trials following other congruents trial (*cC*), $^*p < 0.05$.

order to compare individuals' performance across the additional cognitive control tasks.

Response inhibition was measured by accuracy on No-Go trials that were parametrically manipulated for difficulty. The manipulation involved three different No-Go conditions, which included No-Go trials following one, three, or five Go trials, respectively. Diagnostic group, No-Go trial type, and gender were regressed on accuracy and RT. We found a significant Group × Trial Type interaction [$F_{(2,\ 222)} = 6.54$, $p = 0.002$; **Figure 3A**]. In order to understand this interaction better, we next examined the effects of No-Go condition within each group separately by regressing the No-Go condition on No-Go accuracy for each group. There was a significant effect of No-Go condition on accuracy in TD children, such that when No-Go trials were preceded by increasing numbers of Go trials, TD children had greater accuracy [$F_{(2,\ 96)} = 11.51$, $p < 0.0001$; mean accuracy = 70.5[18.7]%, 77.7[14.7]%, and 81.7[14.0]% for one, three, and five preceding Go trials, respectively]. By contrast, children with 22q11.2DS demonstrated no change in performance across conditions [$F_{(2,\ 126)} = 0.036$, $p = 0.96$; mean accuracy = 72.2[15.9]%, 72.6[15.3]%, and 72.0[17.6]% for one, three, and five preceding Go trials, respectively; Supplementary Table 2 and **Figure 3A**].

In order to examine if the group difference in response inhibition might be due to speed-accuracy trade-offs, RT was measured on consecutive Go trials leading up to a No-Go trial. Diagnostic group, gender, and Go trial number (one through five based on sequential order following a No-Go trial) were regressed on RT. There were no group differences in Go RT [$F_{(1,\ 110)} = 0.22$, $p = 0.64$; Supplementary Table 2 and **Figure 3B**]. Similarly, both groups demonstrated a similar performance pattern, consisting of a relative slowing from the first up to the fourth Go trial following a No-Go trial [$F_{(4,\ 192)} = 30.1$, $p < 0.0001$ for TD; $F_{(4,\ 252)} = 31.2$, $p < 0.0001$ for 22q11.2DS; Supplementary Table 2 and **Figure 3B**]. Thus, while response inhibition was impaired between groups (as measured by No-Go accuracy), this was not due to differences in RT on preceding Go trials.

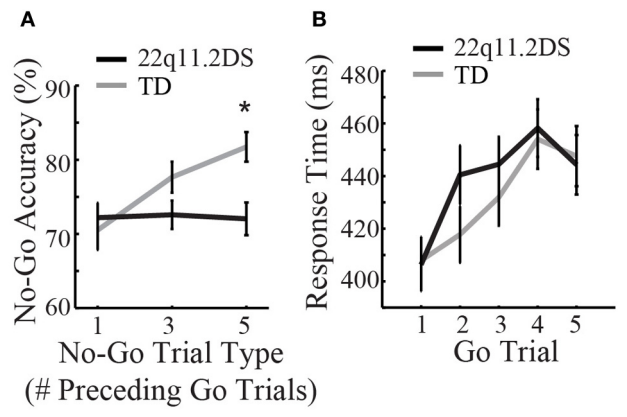


FIGURE 3 | Results of the response inhibition Go/No-Go task. (A) No-Go accuracy differed between groups $^*p < 0.05$, while **(B)** response time on Go trials was similar between groups.

COGNITIVE FLEXIBILITY—VCCS TASK

To examine cognitive flexibility, percent accuracy was regressed against diagnostic group, sorting dimension (predominant or secondary), and gender. There was a significant group difference in accuracy [$F_{(1,\ 109)} = 31.50$, $p < 0.0001$; **Figure 4A**], as well as a significant Group × Dimension interaction, such that children with 22q11.2DS performed more poorly than TD children when sorting by the secondary dimension relative to the predominant dimension [$F_{(1,\ 110)} = 13.41$, $p = 0.0004$]. This was further supported by a significant group difference in the ratio score of accuracy on Dimension 2 divided by that of Dimension 1 [$F_{(1,\ 109)} = 14.45$, $p = 0.0002$; **Figure 4B**]. See Supplementary Table 3 for each group's percent accuracy on both dimensions, as well as the results of statistical tests for group differences in performance on each dimension.

WORKING MEMORY—SOPT TASK

There were two versions of this task (verbal and non-verbal), and each version had three levels of difficulty that from easiest to most difficult involved remembering three, four, or six images, respectively. To examine working memory performance, span and errors were regressed against diagnostic group, level of difficulty, and gender.

For the *verbal* version of the task, there was a significant group difference in span [$F_{(1,\ 113)} = 11.40$, $p = 0.001$], as well as a significant Group × Level interaction, such that children with 22q11.2DS performed more poorly than TD children at higher levels of difficulty relative to lower levels of difficulty [$F_{(2,\ 228)} = 3.39$, $p = 0.04$; **Figure 5B**]. Similarly, there was a significant group difference in number of errors [$F_{(1,\ 113)} = 11.86$, $p = 0.0008$; **Figure 5A**], though the Group × Level interaction here was not quite significant [$F_{(2,\ 228)} = 2.69$, $p = 0.07$]. See

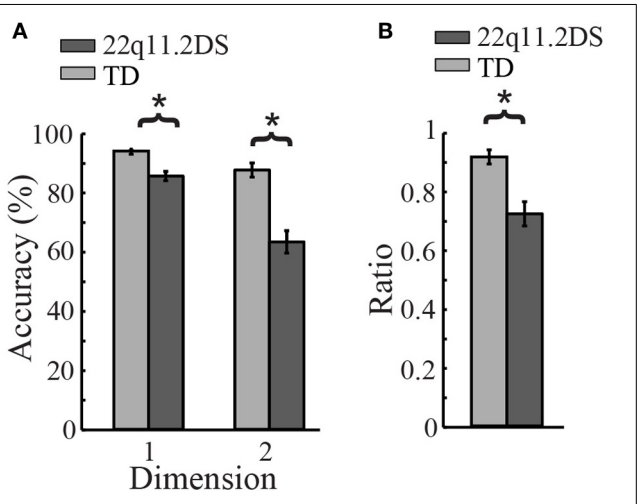


FIGURE 4 | Results of Visually-Cued Card Sort (VCCS), a test of cognitive flexibility. (A) TD children had better accuracy when sorting by both dimensions (predominant and secondary, $^*p < 0.05$). **(B)** Children with 22q11.2DS performed significantly worse when sorting by the secondary dimension relative to the predominant dimension, as indicated by this group difference in the ratio score of accuracy on Dimension 2 divided by that of Dimension 1 ($^*p < 0.05$).

Supplementary Table 4 for group- and level-wise scores on each level, as well as the results of statistical tests for group differences in performance at each level.

For the *non-verbal* version of the task, there was a significant group difference in span [$F_{(1, 100)} = 17.25$, $p = 0.0001$], but no Group × Level interaction [$F_{(2, 202)} = 1.092$, $p = 0.34$; **Figure 5D**]. Similarly, there was a significant group difference in number of errors [$F_{(1, 100)} = 15.08$, $p = 0.0002$] and no Group × Level interaction [$F_{(2,202)} = 0.53$, $p = 0.59$; **Figure 5C**]. The main difference between the results of the verbal vs. the non-verbal version of the test is that there was a Group × Level interaction in performance for the verbal, but not the non-verbal, version of the task. This is likely due to the fact that children with 22q11.2DS performed more poorly than TD children across all levels of the non-verbal version of the task, while they only performed comparably to TD children at easier levels of the verbal version of the task, and worse at more difficult levels. See Supplementary Table 4 for group- and level-wise scores on each level, as well as the results of statistical tests for group differences in performance at each level.

AGE AND COGNITIVE CONTROL

To examine the development of cognitive control in children with 22q11.2DS and TD children, age was included in the within-group regression models. For response inhibition, age was regressed on accuracy for the Stroop and Go/No-Go tasks. We found that age was not related to incongruent accuracy on the Stroop for either group [$F_{(1, 26)} = 0.29$, $p = 0.59$ and $F_{(1, 36)} = 2.52$, $p = 0.12$ for TD and 22q11.2DS, respectively; **Figure 6A**]. The scatterplot of this relationship (**Figure 6A**) illustrates that most TD children are performing at very high levels of accuracy on this task, while variance in performance appears to be increasing in older individuals with 22q11.2DS.

On the Go/No-Go task we found that TD children demonstrated a significant age-related association with No-Go accuracy [$F_{(1, 33)} = 4.91$, $p = 0.03$], such that older TD children performed better on the response inhibition task than younger TD children. By contrast, children with 22q11.2DS demonstrated no relationship of age with No-Go accuracy [$F_{(1, 46)} = 0.53$, $p = 0.47$; **Figure 6B**].

To examine the development of cognitive flexibility in the two groups, age was regressed on percent accuracy for each dimension of the VCCS (Dimensions 1 and 2 for 80 and 20% frequency, respectively), as well as the ratio of Dimension 2 accuracy divided by that of Dimension 1. TD children demonstrated significant associations of age with accuracy on both dimensions [$F_{(1, 47)} = 6.53$, $p = 0.01$ and $F_{(1, 47)} = 11.81$, $p = 0.001$ for Dimension 1 and 2, respectively], as did children with 22q11.2DS [$F_{(1, 59)} = 3.51$, $p = 0.07$ and $F_{(1, 59)} = 5.38$, $p = 0.02$ for Dimension 1 and 2, respectively]. With regard to the ratio of accuracy on Dimension 2 divided by that of Dimension 1, TD children again demonstrated a significant effect of age [$F_{(1, 47)} = 4.48$, $p = 0.04$], while children with 22q11.2DS did not [$F_{(1, 59)} = 2.15$, $p = 0.15$; **Figure 6C**].

To examine the development of working memory, age was regressed against span on the SOPT for each group. Here we found a significant age-related association with span for the verbal version of the SOPT for TD children [$F_{(1, 49)} = 6.11$, $p = 0.02$], as well as children with 22q11.2DS [$F_{(1, 61)} = 6.24$, $p = 0.02$; **Figure 6D**]. Similarly, both groups demonstrated significant age-related associations with span on the non-verbal version of the task [$F_{(1, 44)} = 8.88$, $p = 0.005$ and $F_{(1, 53)} = 20.62$, $p < 0.0001$ for TD and 22q11.2DS, respectively; **Figure 6E**]. See Supplementary Table 5 for a complete list of within-group statistical tests of age on cognitive control outcome measures.

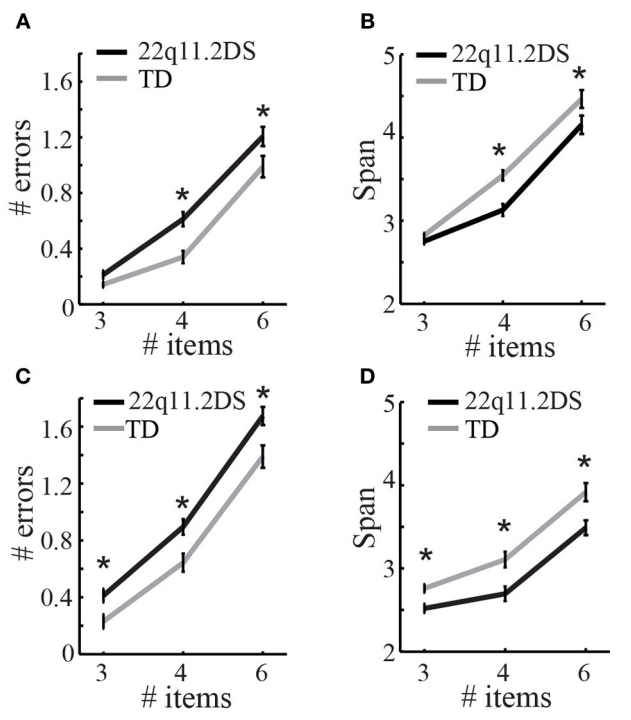


FIGURE 5 | Results of the working memory task, the self-ordered pointing test (SOPT). On the verbal version of the task, children with 22q11.2DS made more errors **(A)** and had a lower span **(B)** on the more difficult trials levels with 4 and 6 items to remember (*$p < 0.05$). On the non-verbal version of the SOPT, children with 22q11.2DS made significantly more errors **(C)** and had a lower span **(D)** across all levels of the non-verbal SOPT, when compared to TD children (*$p < 0.05$).

GENERAL INTELLECTUAL ABILITY AND COGNITIVE CONTROL

There was a significant group difference in full-scale IQ [FSIQ; mean[SD] = 74.8[12.0] for 22q11.2DS and 110.2[12.3] for TD; $F_{(1, 90)} = 200.06$, $p < 0.0001$]. To assess the relationship of general intellectual abilities to cognitive control function, FSIQ was included as a regressor against outcome measures on the cognitive control tasks. On the Stroop task, there were no effects of FSIQ on accuracy on incongruent trials within either of the groups [$F_{(1, 22)} = 0.24$, $p = 0.63$ and $F_{(1, 29)} = 0.79$, $p = 0.38$ for TD and 22q11.2DS, respectively]. Similarly, there were no effects of FSIQ on No-Go accuracy [$F_{(1, 32)} = 0.51$, $p = 0.48$ and $F_{(1, 47)} = 0.88$, $p = 0.35$ for TD and 22q11.2DS, respectively]. On the VCCS test of cognitive flexibility, FSIQ had a significant effect on the Dimension 2/Dimension 1 ratio for TD children [$F_{(1, 33)} = 8.89$, $p = 0.005$] but not children with 22q11.2DS

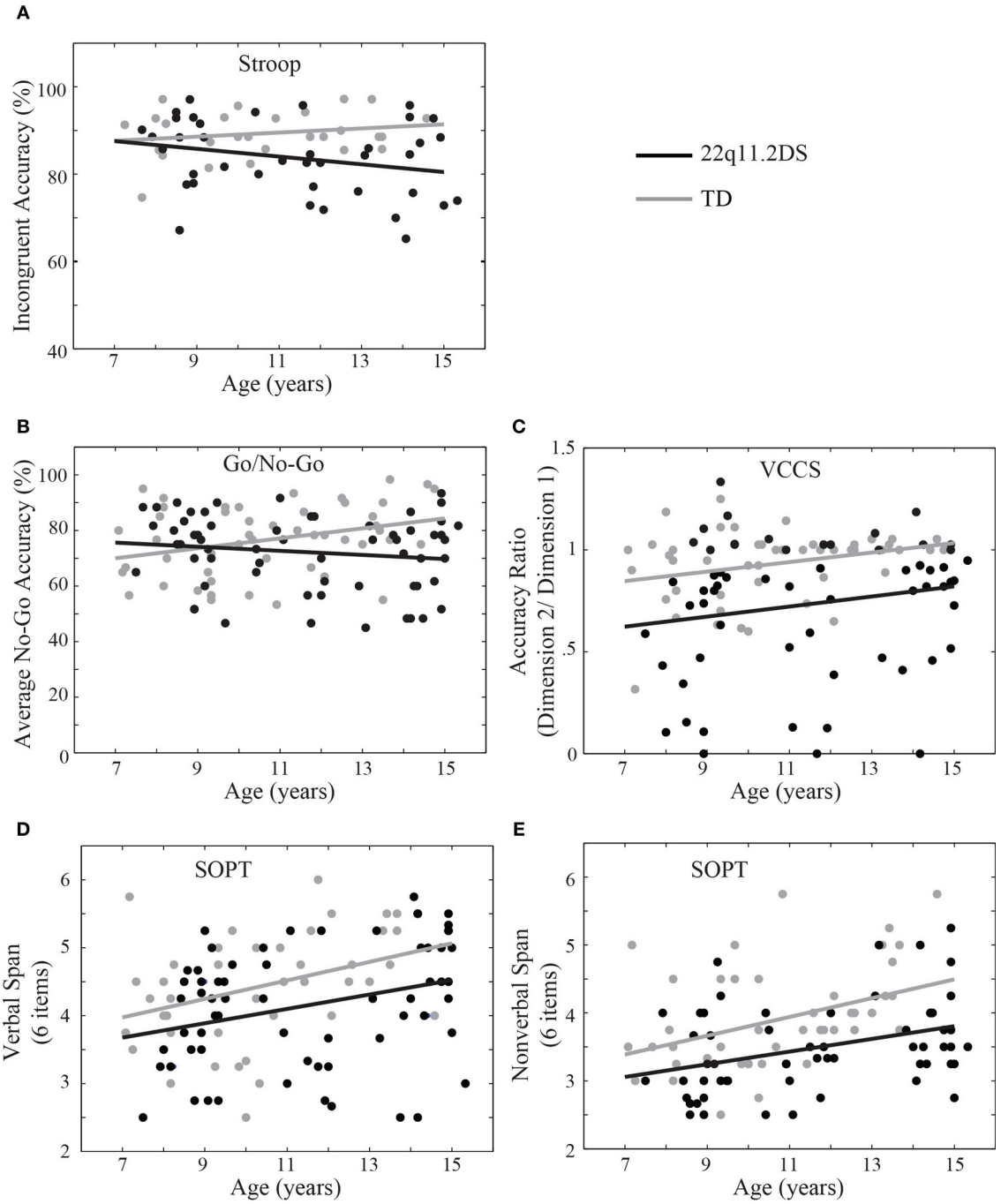


FIGURE 6 | Age-related associations with cognitive control component processes. Age was not related to incongruent accuracy on the Stroop for either group **(A)**. On the Go/No-Go task, TD children demonstrated a significant age-related association with No-Go accuracy ($p < 0.05$) while children with 22q11.2DS did not **(B)**. Similarly, on the Visually-Cued Card Sort (VCCS), age was significantly associated with the ratio of Dimension 2 accuracy divided by that of Dimension 1 for TD children ($p < 0.05$) but not those with 22q11.2DS **(C)**. Age significantly correlated with span on the Self-Ordered Pointing Test (SOPT) for both TD and 22q11.2DS children on the verbal **(D)** and nonverbal **(E)** versions of the task ($p < 0.05$).

$[F_{(1,\ 44)} = 2.43,\ p = 0.13]$. On the SOPT test of working memory, the only significant within-group relationship of FSIQ with span was that of non-verbal span with FSIQ in TD children $[F_{(1,\ 30)} = 9.91, p = 0.004]$.

COMT AND COGNITIVE CONTROL

First, we wanted to visualize the relationship of COMT genotype to performance on the different cognitive control tasks, in order to assess whether or not specific COMT genotypes might

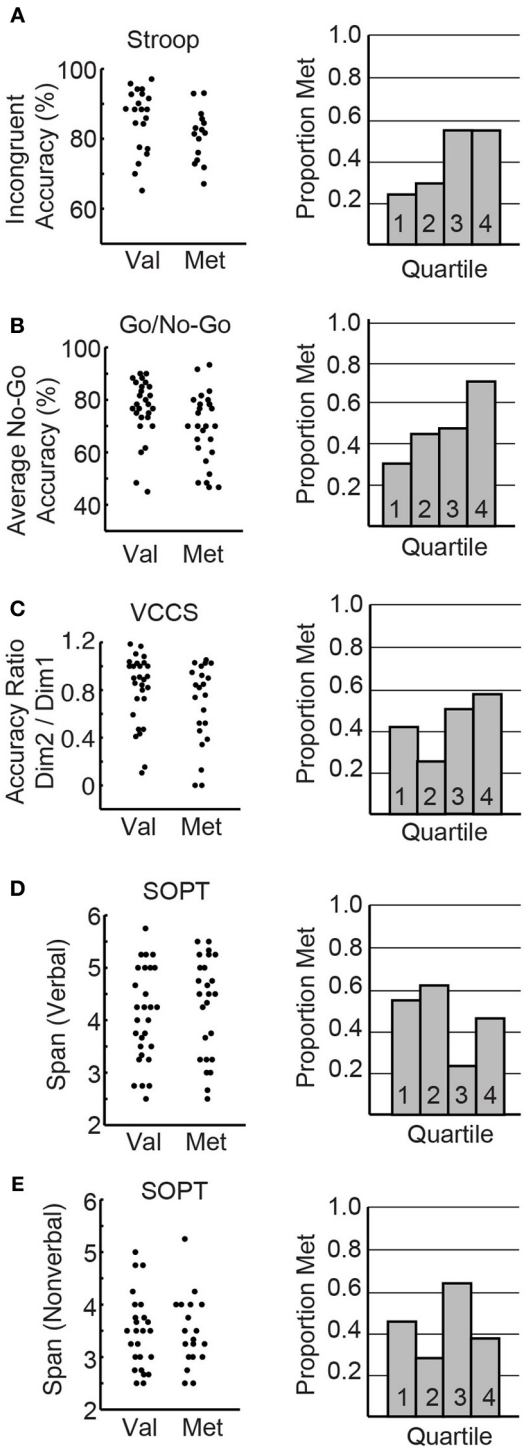


FIGURE 7 | COMT and cognitive control. The left panel depicts the primary outcome measures for each task graphed as a function of COMT variant for the individual children with 22q11.2DS. **(A)** Incongruent accuracy on the Stroop task. **(B)** Average No-Go accuracy on the response inhibition Go/No-Go task. **(C)** Accuracy ratio (Dimension 2/Dimension 1) on the VCCS. **(D,E)** Verbal and non-verbal span, respectively, on the most difficult level of the self-ordered pointing test (6 items to remember). The right panels of the figure depict the proportion of individuals within each performance quartile of the particular task that had the Met variant of the COMT gene.

account for some of the variance that is seen among individuals with 22q11.2DS. For each task, we graphed the primary outcome measure as a function of genotype for the children with 22q11.2DS (**Figures 7A–E**). Qualitatively, it appeared that on the response inhibition tasks, there were more individuals with the Met allele performing poorly relative to those with the Val allele. In order to quantify this observation, we split the participants into four groups based on their performance. The first group included those performing in the top quartile of the sample (highest performers), down to the fourth group that consisted of those performing in the fourth quartile of the sample (lowest performers). We then graphed the proportion of individuals within each quartile that had the Met allele (calculated by taking the number of participants within that sample that had the Met allele, divided by the total number of participants in that quartile; **Figures 7A–E**).

We also assessed potential COMT effects using regression models. COMT genotype was included as a regressor against outcome measures on the cognitive control tasks in children with 22q11.2DS. On the Stroop task, there was no effect of COMT on incongruent accuracy [$F_{(1,\ 32)} = 1.40$, $p = 0.26$]. By contrast, on the Go/No-Go task, COMT genotype had a significant effect on overall No-Go accuracy [$F_{(1,\ 50)} = 4.54$, $p = 0.04$], such that individuals with the Met allele had lower accuracy. There was no effect of COMT on the ratio of Dimension 2/Dimension 1 accuracy on the VCCS task [$F_{(1,\ 46)} = 1.99$, $p = 0.16$]. Similarly there was no effect of COMT on working memory span for either version of the SOPT task [$F_{(1,\ 49)} = 0.89$, $p = 0.35$ and $F_{(1,\ 41)} = 0.40$, $p = 0.53$ for verbal and non-verbal, respectively].

DISCUSSION

The present study was designed to investigate cognitive control and its age-related development in a cross-sectional sample of children with 22q11.2DS. As expected based on the literature (Sobin et al., 2005; Lewandowski et al., 2007; Campbell et al., 2010), when compared to TD controls, children with 22q11.2DS were impaired on all three cognitive control foundational processes: response inhibition, cognitive flexibility, and working memory. The advantage of this study is that it enabled us to examine individual performance patterns across a battery of tasks within the same sample of participants, thus identifying relative strengths and weaknesses in cognitive control component processes that might generate hypotheses about specific mechanisms underpinning cognitive control impairments. Importantly, by examining these processes across an age range of children with 22q11.2DS and TD controls, we were able to conduct a cross-sectional analysis of developmental trajectories.

As expected, TD children demonstrated a significant effect of age on most cognitive control component processes, such that older children had better performance relative to their younger counterparts. The only measure on which TD children did not demonstrate an age-related association was that of Stroop incongruent accuracy, likely due to nearly ceiling effects across all ages (**Figure 6A**). By contrast, children with 22q11.2DS demonstrated no age-related associations within our 7–14 year age range on four of the tasks, including Stroop, Go/No-Go, and VCCS. Analysis of individual performance patterns on the response inhibition tasks

(Stroop and Go/No-Go) suggested that some of the older children with 22q11.2DS performed similarly to TD children while others performed much worse. Thus, an atypical developmental trajectory of response inhibition in this population was due to increased variability of performance in older individuals with the disorder. The inter-individual variability seen in older individuals with 22q11.2DS may contain great value with respect to identifying individuals whose inhibitory function is developing atypically relative to their peers, thus providing insight into mechanisms that might be underpinning variability within the group. Distinguishing measures such as these, especially those that have been linked to cognitive dysfunction in schizophrenia, are valuable targets to explore for better understanding individuals that are at greater risk for psychopathology. It will be important to explore these developmental patterns longitudinally in future studies.

This same sample of participants also demonstrated an atypical age-related association with cognitive flexibility. While general card sorting ability on the task had a similar age-related association in the two groups, the ability to sort by the less dominant dimension was not only impaired in children with 22q11.2DS, but also did not show the typical effect of improving with age that was apparent in the TD participants. Similar to the response inhibition results, this overall group effect was due to increased variability of performance in older individuals with 22q11.2DS, with some performing well and others highly impaired.

In contrast to the atypical age effects seen in response inhibition and cognitive flexibility, the children with 22q11.2DS demonstrated a typical relationship of age with span on the working memory task. This preliminary cross-sectional sample suggests that, despite an overall impairment in performance on this task, the development of this component of cognitive control in 22q11.2DS might be more typical than the others. One possible implication here is that the neural circuitry supporting working memory is developing and becoming more efficient at a rate similar to TD individuals. Alternatively, it is possible that compensatory mechanisms support improved performance on the working memory task in this age range of individuals with 22q11.2DS.

It is important to think about these results in the context of a framework for cognitive control, while remembering that the distinct sub-components are neither pure nor perfect with respect to their distinctions, as well as the tasks proposed to measure them. As described by Miyake et al. (2000), this system is likely composed of foundational cognitive control components that are both distinct and interrelated. Additionally, with respect to the tasks designed to measure these components, there will surely be overlap in the functions required for completing each task. The Go/No-Go task requires inhibitory control in order to inhibit a pre-potent response to press the button on frequently occurring "Go" trials. This task also requires some working memory in order to remember which stimuli are indicative of a Go trial and which stimuli represent a No-Go trial. The VCCS requires participants to follow specific rules and to be cognitively flexible in order to respond appropriately to the given rules that change according to certain criteria. Working memory is necessary to remember the current rules at hand. Additionally, inhibitory control is required in order to inhibit the inclination to respond according to the predominant sorting dimension. Finally, the SOPT requires participants to hold a number of items in working memory, while also comparing responses that have already been made with those that will be made in the future (self-monitoring). This type of behavior also requires some degree of planning and organization. While it is important to recognize that the neurocognitive tasks here might be multi-componential to some degree, their unique emphasis on specific cognitive processes is important to recognize, and the overlapping nature provides an opportunity to compare performance on different components with respect to their primary and overlapping functionalities.

The results of the current study suggest a specific aberration in 22q11.2DS in the development of networks mediating response inhibition and cognitive flexibility. One unifying feature of the response inhibition and cognitive flexibility tasks that distinguishes them from the working memory task is that the former two both require the ability to inhibit a pre-potent response. Given the component of inhibitory control that is required to sort by the less dominant dimension on the VCCS, it is unclear the extent to which difficulties with response inhibition might underlie performance on this task of cognitive flexibility. One approach to examining the specific and interrelated nature of cognitive control component processes could be through latent variable analyses and computational modeling (Miyake et al., 2000; Friedman and Miyake, 2004; Miyake and Friedman, 2012). These would be important studies in the future for better understanding the most specific nature of cognitive control impairments in 22q11.2DS.

Cognitive control impairments are exceedingly common in other neurodevelopmental and psychiatric disorders. In order to examine the specificity of cognitive control impairments in 22q11.2DS, and whether or not group differences in global cognitive functioning (non-specific to 22q11.2DS) were driving the results, FSIQ was included as a regressor against outcome measures on the cognitive control tasks. We found that FSIQ was not related to task performance in the children with 22q11.2DS, thus suggesting that the observed impairments in cognitive control were not being driven by global cognitive functioning, as measured by FSIQ. It is important to mention, however, that FSIQ alone is not necessarily a comprehensive measure of global cognitive functioning, and that future work will be needed in order to more directly examine the relationship of general intelligence to cognitive control in 22q11.2DS. For example, it would be important to examine the relationship of cognitive control impairments in 22q11.2DS to fluid intelligence, which is believed to reflect abstract reasoning and problem solving skills, a functionality that is impaired after lesions of the frontal lobe (Duncan et al., 1995). It has been demonstrated that, in a population of patients with frontal lesions, there were no specific deficits related to cognitive control once fluid intelligence was taken into account (Roca et al., 2010). It is likely, however, that cognitive control impairments are not fully explained by fluid intelligence. Other evidence suggests that the different cognitive control component processes (inhibition, cognitive flexibility, and working memory) are differentially related to fluid intelligence (Friedman et al., 2006). The specificity of these impairments and their relationship to fluid intelligence remain to be parsed in developmental disorders, and this will be an important question to pursue in children with 22q11.2DS.

Another important issue related to understanding the specificity of cognitive impairments and their developmental courses in 22q11.2DS is the selection of an appropriate comparison population. In the present study there are limitations associated with matching developmentally delayed individuals with age-matched typical controls. However, matching by mental age or by cognitive ability would introduce additional variables and confounds. More specifically, a comparison group matched by cognitive ability would involve a highly heterogeneous sample of participants with many different etiologies. At present we felt that, even despite the limitations, it was important to observe performance in 22q11.2DS relative to age-matched TD controls. For one, it affords the opportunity to draw comparisons from a representative control sample as opposed to a heterogenous sample (Dennis et al., 2009). Importantly, this design allows us to estimate the magnitude of impairment in 22q11.2DS relative to age-matched TD controls, thus establishing a baseline that can be used as a reference in the future for studies of intervention. Potential timelines for cognitive control development have been described in TD children (Huizinga et al., 2006; Best and Miller, 2010). In the current study we wanted to assess how the developmental trajectories of cognitive control in 22q11.2DS might compare to the standard in TD individuals, given that atypical neurodevelopmental trajectories are exceedingly common in childhood psychiatric disorders (Shaw et al., 2010). For this comparison, it is important to include TD participants in order to first replicate the existing data and show that the current tasks are validly reproducing well-established developmental time courses in the TD group. Subsequently, we can accurately assess the differences in developmental trajectories between TD and 22q11.2DS, as well as within children with 22q11.2DS.

In addition to cognitive analysis, the current study also examined COMT genotype as a function of performance on the different cognitive control tasks, in order to assess whether or not specific COMT genotypes might account for some of the variance that is seen among individuals with 22q11.2DS. Interestingly, it appeared that the children with 22q11.2DS who were hemizygous for the Met variant of the COMT gene performed more poorly on the tasks of response inhibition relative to their peers with 22q11.2DS who were hemizygous for the Val allele. Though this relationship was only statistically significant when assessing performance on the Go/No-Go task, it appeared that there was a trend toward this relationship in the other tasks with inhibitory requirements, including the Stroop and VCCS (see **Figures 7A–C**). By contrast, this was not the case for the SOPT task of working memory (**Figures 7D,E**). This is an interesting dissociation, given that previous studies have suggested that inhibitory tasks are more dependent on DA than the SOPT (Diamond et al., 1997; Collins et al., 1998). These results indicate that participants with 22q11.2DS who were hemizygous for the Met allele tended to perform worse on the tasks that have previously been suggested to be DA-dependent. According to the model that the effect of DA on cognition follows an inverse U pattern, with an optimal range of DA involving not too much or too little of the neurotransmitter, it is reasonable to assume that children with 22q11.2DS and the Met allele for COMT are at a disadvantage relative to those with Val. After all, since children with 22q11.2DS are already hemizygous for COMT, it is likely that they have less prefrontal COMT activity and higher levels of DA. Thus, hemizygosity for Val, the variant with greater catalytic activity, would be more advantageous for maintaining a position closer to the optimal peak of prefrontal DA as it relates to higher-level cognitive processes. This hypothesis will have to be tested further in the future.

It is not surprising that discrepancies in COMT effects in 22q11.2DS are often reported in the literature. After all, the effects of a single gene are not likely to be very powerful, and impact might also vary as a function of other factors such as age, gender (Kates et al., 2006), or other genetic variants (Vorstman et al., 2009). In addition to the issue of power, another reason for the limited and inconclusive reports on the relationship of genetic variants to cognitive function is that genetics are likely influenced by environmental factors. Two noteworthy factors are stress and anxiety. While the genetics of 22q11.2DS predispose individuals to susceptibility for greater stress and anxiety, it is possible that mechanisms for coping and reducing these influences will contribute to better adaptive function and thus better long-term outcomes (Beaton and Simon, 2011; Angkustsiri et al., 2012).

It is reasonable to assume that the observed cognitive control impairments in 22q11.2DS are in some way mediated by the genetics of the disorder, and are subserved by underlying impairments in neural architecture that supports these cognitive processes. Cognitive control is largely mediated by activity within the prefrontal cortex (PFC) and reciprocal connections between the PFC and subcortical networks. In humans and monkeys, damage to the dorsolateral PFC (dlPFC) impairs performance on the Go/No-Go task (Iversen and Mishkin, 1970), the VCCS (Passingham, 1972; Dias et al., 1996), and the SOPT (Petrides and Milner, 1982; Petrides, 1991). Additionally, neuroimaging studies have demonstrated that the dlPFC is more active during each respective cognitive task when compared to a control task (Petrides et al., 1993; Berman et al., 1995; Casey et al., 1997).

Frontally-mediated regulation of cognitive control is often modulated by subcortical circuitry. One of the key components of this system is frontostriatal circuitry, which involves neuronal loops connecting the PFC, thalamus, and basal ganglia. The basal ganglia consist of interconnected subcortical nuclei that receive major input from the cerebral cortex and thalamus, and then connect back to the cerebral cortex via the thalamus (Alexander et al., 1986). There is some evidence that these circuits are atypical in 22q11.2DS. Structural imaging studies have demonstrated GM reduction and dysfunction in 22q11.2DS (Shashi et al., 2010), as well as alterations in midline cortical thickness and gyrification patterns (Bearden et al., 2009). There is also evidence for atypical basal ganglia structure in 22q11.2DS (Sugama et al., 2000; Eliez et al., 2002), as well as atypical structural connectivity within frontal networks (Simon et al., 2008). Functional imaging studies have also demonstrated irregularities in these networks in children with 22q11.2DS when compared to TD children, including atypical parietal activity during a Go/No-Go task (Gothelf et al., 2007a), as well as hypoactivation of dorsolateral PFC during performance on a working memory task (Kates et al., 2007).

With respect to the structural and functional developmental of cognitive control neural networks in 22q11.2DS, evidence suggests that the developmental trajectory of cortical gyrification is atypical in children with 22q11.2DS relative to TD children in this age range (6–15 years) (Srivastava et al., 2011). The specific nature and timing of these trajectories are still unclear, however, and to date there have only been a few longitudinal studies of developmental trajectories of brain structure in 22q11.2DS (Gothelf et al., 2007b; Schaer et al., 2009; Kates et al., 2011; Kunwar et al., 2012). While these studies indicated neuroanatomical differences in frontal and parietal regions in children and adolescents with 22q11.2DS relative to TD individuals, evidence for atypical development trajectories was inconsistent. Larger samples of longitudinal studies during this critical developmental time period will be important for more directly examining the development of brain and behavior relationships responsible for cognitive control in 22q11.2DS.

A better understanding of genes, brain, behavior, and external modulatory components of cognitive control in 22q11.2DS is most relevant given the high risk of schizophrenia in this population. Approximately 25% of individuals with 22q11.2DS will develop schizophrenia by adulthood (Murphy et al., 1999), rendering it the highest genetic risk factor for the disorder after having two parents or a twin sibling with schizophrenia. There is evidence for attenuated cognitive control impairments among first-degree relatives of individuals with schizophrenia, suggesting that these deficits might be part of an endophenotype related to genetic susceptibility for the disorder (Snitz et al., 2006). Thus, the results of the current study pose interesting questions as to whether aberrant response inhibition might be part of an endophenotype for schizophrenia risk in 22q11.2DS, and if so might the lower-performing older individuals with 22q11.2DS be the individuals at greatest risk for conversion? These are questions that will be explored in the future via longitudinal analyses and correlations with measures of psychosis. In this manner, we will be able to directly examine the potential for these kinds of tasks as non-invasive diagnostic measures for risk probability, or as evaluative tools for the efficacy of targeted interventions (Carter and Barch, 2007).

In sum, these results point toward a specific aberration in the development of systems mediating response inhibition in a subset of the children with 22q11.2DS, at a critical age when these individuals are at significant risk for developing schizophrenia. Though the present study was cross-sectional in design, it provides a valuable starting point for longitudinal analyses. In the future it will be important to directly examine developmental trajectories that integrate genetic, physiological, neurocognitive, and clinical psychosis measures in order to obtain a most comprehensive picture of modulatory factors pertaining to the development of cognitive control, as well as clinical and adaptive outcomes.

ACKNOWLEDGMENTS

We would like to thank all the families that participated in our research. Funding for the current study was made possible by NIH grants R01HD02974 (to Tony J. Simon) and UL1 RR024146 from the National Center for Medical Research. Furthermore, the first author was supported by a Training Grant from the National Institute on Deafness and Other Communication Disorders (5T32DC008072). The funding bodies had no further role in the study design; in the acquisition, analysis, and interpretation of data; in the writing of the manuscript; and in the decision to submit the paper for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00566/abstract

REFERENCES

Alexander, G. E., DeLong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381. doi: 10.1146/annurev.ne.09.030186.002041

Angkustsiri, K., Leckliter, I., Tartaglia, N., Beaton, E. A., Enriquez, J., and Simon, T. J. (2012). An examination of the relationship of anxiety and intelligence to adaptive functioning in children with chromosome 22q11.2 deletion syndrome. *J. Dev. Behav. Pediatr.* 33, 713–720. doi: 10.1097/DBP.0b013e318272dd24

Baker, K., Baldeweg, T., Sivagnanasundaram, S., Scambler, P., and Skuse, D. (2005). COMT Val108/158 Met modifies mismatch negativity and cognitive function in 22q11 deletion syndrome. *Biol. Psychiatry* 58, 23–31. doi: 10.1016/j.biopsych.2005.03.020

Bassett, A. S., Chow, E. W. C., AbdelMalik, P., Gheorghiu, M., Husted, J., and Weksberg, R. (2003). The schizophrenia phenotype in 22q11 deletion syndrome. *Am. J. Psychiatry* 160, 1580–1586. doi: 10.1176/appi.ajp.160.9.1580

Bearden, C. E., Jawad, A. F., Lynch, D. R., Sokol, S., Kanes, S. J., McDonald-McGinn, D. M., et al. (2004). Effects of a functional COMT polymorphism on prefrontal cognitive function in patients with 22q11.2 deletion syndrome. *Am. J. Psychiatry* 161, 1700–1702. doi: 10.1176/appi.ajp.161.9.1700

Bearden, C. E., van Erp, T. G. M., Dutton, R. A., Lee, A. D., Simon, T. J., Cannon, T. D., et al. (2009). Alterations in midline cortical thickness and gyrification patterns mapped in children with 22q11.2 deletions. *Cereb. Cortex* 19, 115–126. doi: 10.1093/cercor/bhn064

Beaton, E. A., and Simon, T. J. (2011). How might stress contribute to increased risk for schizophrenia in children with chromosome 22q11.2 deletion syndrome? *J. Neurodev. Disord.* 3, 68–75. doi: 10.1007/s11689-010-9069-9

Berman, K. F., Ostrem, J. L., Randolph, C., Gold, J., Goldberg, T. E., Coppola, R., et al. (1995). Physiological activation of a cortical network during performance of the Wisconsin Card Sorting Test: a positron emission tomography study. *Neuropsychologia* 33, 1027–1046. doi: 10.1016/0028-3932(95)00035-2

Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x

Bish, J. P., Ferrante, S. M., McDonald-McGinn, D., Zackai, E., and Simon, T. J. (2005). Maladaptive conflict monitoring as evidence for executive dysfunction in children with chromosome 22q11.2 deletion syndrome. *Dev. Sci.* 8, 36–43. doi: 10.1111/j.1467-7687.2005.00391.x

Brewer, W. J., Francey, S. M., Wood, S. J., Jackson, H. J., Pantelis, C., Phillips, L. J., et al. (2005). Memory impairments identified in people at ultra-high risk for psychosis who later develop first-episode psychosis. *Am. J. Psychiatry* 162, 71–78. doi: 10.1176/appi.ajp.162.1.71

Campbell, L. E., Azuma, R., Ambery, F., Stevens, A., Smith, A., Morris, R. G., et al. (2010). Executive functions and memory abilities in children with 22q11.2 deletion syndrome. *Aust. N. Z. J. Psychiatry* 44, 364–371. doi: 10.3109/00048670903489882

Cannon, T. D., van Erp, T. G. M., Bearden, C. E., Loewy, R., Thompson, P., Toga, A. W., et al. (2003). Early and late neurodevelopmental influences in the prodrome to schizophrenia: contributions of genes, environment, and their interactions. *Schizophr. Bull.* 29, 653–669. doi: 10.1093/oxfordjournals.schbul.a007037

Carlson, C., Sirotkin, H., Pandita, R., Goldberg, R., McKie, J., Wadey, R., et al. (1997). Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients. *Am. J. Hum. Genet.* 61, 620–629. doi: 10.1086/515508

Carter, C. S., and Barch, D. M. (2007). Cognitive neuroscience-based approaches to measuring and improving treatment effects on cognition

in schizophrenia: the CNTRICS initiative. *Schizophr. Bull.* 33, 1131–1137. doi: 10.1093/schbul/sbm081

Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N., et al. (1997). A developmental functional MRI study of prefrontal activation during performance of a Go-No-Go task. *J. Cogn. Neurosci.* 9, 835–847. doi: 10.1162/jocn.1997.9.6.835

Collins, P., Roberts, A. C., Dias, R., Everitt, B. J., and Robbins, T. W. (1998). Perseveration and strategy in a novel spatial self-ordered sequencing task for nonhuman primates: effects of excitotoxic lesions and dopamine depletions of the prefrontal cortex. *J. Cogn. Neurosci.* 10, 332–354. doi: 10.1162/089892998562771

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., and Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *J. Int. Neuropsychol. Soc.* 15, 331–343. doi: 10.1017/S1355617709090481

Diamond, A., Prevor, M. B., Callender, G., and Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monogr. Soc. Res. Child Dev.* 62, i–v, 1–208. doi: 10.2307/1166208

Dias, R., Robbins, T. W., and Roberts, A. C. (1996). Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behav. Neurosci.* 110, 872–886. doi: 10.1037/0735-7044.110.5.872

Duncan, J., Burgess, P., and Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia* 33, 261–268. doi: 10.1016/0028-3932(94)00124-8

Eliez, S., Barnea-Goraly, N., Schmitt, J. E., Liu, Y., and Reiss, A. L. (2002). Increased basal ganglia volumes in velo-cardio-facial syndrome (deletion 22q11.2). *Biol. Psychiatry* 52, 68–70. doi: 10.1016/S0006-3223(02)01361-6

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135. doi: 10.1037/0096-3445.133.1.101

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., and Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychol. Sci.* 17, 172–179. doi: 10.1111/j.1467-9280.2006.01681.x

Glaser, B., Debbané, M., Hinard, C., Morris, M. A., Dahoun, S. P., Antonarakis, S. E., et al. (2006). No evidence for an effect of COMT Val158Met genotype on executive function in patients with 22q11 deletion syndrome. *Am. J. Psychiatry* 163, 537–539. doi: 10.1176/appi.ajp.163.3.537

Gothelf, D., Hoeft, F., Hinard, C., Hallmayer, J. F., Stoecker, J. V. D., Antonarakis, S. E., et al. (2007a). Abnormal cortical activation during response inhibition in 22q11.2 deletion syndrome. *Hum. Brain Mapp.* 28, 533–542. doi: 10.1002/hbm.20405

Gothelf, D., Penniman, L., Gu, E., Eliez, S., and Reiss, A. L. (2007b). Developmental trajectories of brain structure in adolescents with 22q11.2 deletion syndrome: a longitudinal study. *Schizophr. Res.* 96, 72–81. doi: 10.1016/j.schres.2007.07.021

Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Iversen, S. D., and Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp. Brain Res.* 11, 376–386. doi: 10.1007/BF00237911

Kates, W. R., Antshel, K. M., Abdulsabur, N., Colgan, D., Funke, B., Fremont, W., et al. (2006). A gender-moderated effect of a functional COMT polymorphism on prefrontal brain morphology and function in velo-cardio-facial syndrome (22q11.2 deletion syndrome). *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 141B, 274–280. doi: 10.1002/ajmg.b.30284

Kates, W. R., Antshel, K. M., Faraone, S. V., Fremont, W. P., Higgins, A. M., Shprintzen, R. J., et al. (2011). Neuroanatomic predictors to prodromal psychosis in velocardiofacial syndrome (22q11.2 deletion syndrome): a longitudinal study. *Biol. Psychiatry* 69, 945–952. doi: 10.1016/j.biopsych.2010.10.027

Kates, W. R., Krauss, B. R., Abdulsabur, N., Colgan, D., Antshel, K. M., Higgins, A. M., et al. (2007). The neural correlates of non-spatial working memory in velocardiofacial syndrome (22q11.2 deletion syndrome). *Neuropsychologia* 45, 2863–2873. doi: 10.1016/j.neuropsychologia.2007.05.007

Kimberg, D. Y., and D'Esposito, M. (2003). Cognitive effects of the dopamine receptor agonist pergolide. *Neuropsychologia* 41, 1020–1027. doi: 10.1016/S0028-3932(02)00317-2

Kunwar, A., Ramanathan, S., Nelson, J., Antshel, K. M., Fremont, W., Higgins, A. M., et al. (2012). Cortical gyrification in velo-cardio-facial (22q11.2 deletion) syndrome: a longitudinal study. *Schizophr. Res.* 137, 20–25. doi: 10.1016/j.schres.2012.01.032

Lencz, T., Smith, C. W., McLaughlin, D., Auther, A., Nakayama, E., Hovey, L., et al. (2006). Generalized and specific neurocognitive deficits in prodromal schizophrenia. *Biol. Psychiatry* 59, 863–871. doi: 10.1016/j.biopsych.2005.09.005

Lewandowski, K. E., Shashi, V., Berry, P. M., and Kwapil, T. R. (2007). Schizophrenic-like neurocognitive deficits in children and adolescents with 22q11 deletion syndrome. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 144B, 27–36. doi: 10.1002/ajmg.b.30379

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Murphy, K. C., Jones, L. A., and Owen, M. J. (1999). High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch. Gen. Psychiatry* 56, 940–945. doi: 10.1001/archpsyc.56.10.940

Oskarsdóttir, S., Vujic, M., and Fasth, A. (2004). Incidence and prevalence of the 22q11 deletion syndrome: a population-based study in Western Sweden. *Arch. Dis. Child.* 89, 148–151. doi: 10.1136/adc.2003.026880

Passingham, R. E. (1972). Non-reversal shifts after selective prefrontal ablations in monkeys (*Macaca mulatta*). *Neuropsychologia* 10, 41–46. doi: 10.1016/0028-3932(72)90041-3

Petrides, M. (1991). Functional specialization within the dorsolateral frontal cortex for serial order memory. *Proc. Biol. Sci.* 246, 299–306. doi: 10.1098/rspb.1991.0158

Petrides, M., Alivisatos, B., Evans, A. C., and Meyer, E. (1993). Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. *Proc. Natl. Acad. Sci. U.S.A.* 90, 873–877. doi: 10.1073/pnas.90.3.873

Petrides, M., and Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia* 20, 249–262. doi: 10.1016/0028-3932(82)90100-2

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., et al. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain* 133, 234–247. doi: 10.1093/brain/awp269

Scambler, P. J. (2000). The 22q11 deletion syndromes. *Hum. Mol. Genet.* 9, 2421–2426. doi: 10.1093/hmg/9.16.2421

Schaer, M., Debbané, M., Bach Cuadra, M., Ottet, M.-C., Glaser, B., Thiran, J.-P., et al. (2009). Deviant trajectories of cortical maturation in 22q11.2 deletion syndrome (22q11DS): a cross-sectional and longitudinal study. *Schizophr. Res.* 115, 182–190. doi: 10.1016/j.schres.2009.09.016

Shapiro, H. M., Takarae, Y., Harvey, D., Cabaral, M., and Simon, T. J. (2012). A cross-sectional study of the development of volitional control of spatial attention in children with chromosome 22q11.2 deletion syndrome. *J. Neurodev. Disord.* 4:5. doi: 10.1186/1866-1955-4-5

Shapiro, H. M., Wong, L. M., and Simon, T. J. (2013). A cross-sectional analysis of the development of response inhibition in children with chromosome 22q11.2 deletion syndrome. *Front. Psychiatry* 4:81. doi: 10.3389/fpsyt.2013.00081

Shashi, V., Keshavan, M. S., Howard, T. D., Berry, M. N., Basehore, M. J., Lewandowski, E., et al. (2006). Cognitive correlates of a functional COMT polymorphism in children with 22q11.2 deletion syndrome. *Clin. Genet.* 69, 234–238. doi: 10.1111/j.1399-0004.2006.00569.x

Shashi, V., Kwapil, T. R., Kaczorowski, J., Berry, M. N., Santos, C. S., Howard, T. D., et al. (2010). Evidence of gray matter reduction and dysfunction in chromosome 22q11.2 deletion syndrome. *Psychiatry Res.* 181, 1–8. doi: 10.1016/j.pscychresns.2009.07.003

Shaw, P., Gogtay, N., and Rapoport, J. (2010). Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. *Hum. Brain Mapp.* 31, 917–925. doi: 10.1002/hbm.21028

Shprintzen, R. J. (2008). Velo-cardio-facial syndrome: 30 Years of study. *Dev. Disabil. Res. Rev.* 14, 3–10. doi: 10.1002/ddrr.2

Simon, T. J. (2008). A new account of the neurocognitive foundations of impairments in space, time and number processing in children with chromosome 22q11.2 deletion syndrome. *Dev. Disabil. Res. Rev.* 14, 52–58. doi: 10.1002/ddrr.8

Simon, T. J., Bish, J. P., Bearden, C. E., Ding, L., Ferrante, S., Nguyen, V., et al. (2005). A multilevel analysis of cognitive dysfunction and psychopathology

associated with chromosome 22q11.2 deletion syndrome in children. *Dev. Psychopathol.* 17, 753–784. doi: 10.1017/S0954579405050364

Simon, T. J., and Luck, S. J. (2011). *Attentional Impairments in Children with Chromosome 22q11.2 Deletion Syndrome, 2nd Edn*. New York, NY: Guilford Press.

Simon, T. J., Wu, Z., Avants, B., Zhang, H., Gee, J. C., and Stebbins, G. T. (2008). Atypical cortical connectivity and visuospatial cognitive impairments are related in children with chromosome 22q11.2 deletion syndrome. *Behav. Brain Funct.* 4:25. doi: 10.1186/1744-9081-4-25

Snitz, B. E., MacDonald, A. W., and Carter, C. S. (2006). Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: a meta-analytic review of putative endophenotypes. *Schizophr. Bull.* 32, 179–194. doi: 10.1093/schbul/sbi048

Sobin, C., Kiley-Brabeck, K., and Karayiorgou, M. (2005). Lower prepulse inhibition in children with the 22q11 deletion syndrome. *Am. J. Psychiatry* 162, 1090–1099. doi: 10.1176/appi.ajp.162.6.1090

Srivastava, S., Buonocore, M. H., and Simon, T. J. (2011). Atypical developmental trajectory of functionally significant cortical areas in children with chromosome 22q11.2 deletion syndrome. *Hum. Brain Mapp.* 33, 213–223. doi: 10.1002/hbm.21206

Stoddard, J., Beckett, L., and Simon, T. J. (2011). Atypical development of the executive attention network in children with chromosome 22q11.2 deletion syndrome. *J. Neurodev. Disord.* 3, 76–85. doi: 10.1007/s11689-010-9070-3

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol. Gen.* 6, 643–666. doi: 10.1037/h0054651

Sugama, S., Bingham, P. M., Wang, P. P., Moss, E. M., Kobayashi, H., and Eto, Y. (2000). Morphometry of the head of the caudate nucleus in patients with velocardiofacial syndrome (del 22q11.2). *Acta Paediatr.* 89, 546–549. doi: 10.1080/080352500750027826

Takarae, Y., Schmidt, L., Tassone, F., and Simon, T. J. (2009). Catechol-O-methyltransferase polymorphism modulates cognitive control in children with chromosome 22q11.2 deletion syndrome. *Cogn. Affect. Behav. Neurosci.* 9, 83–90. doi: 10.3758/CABN.9.1.83

Vorstman, J. A. S., Turetsky, B. I., Sijmens-Morcus, M. E. J., de Sain, M. G., Dorland, B., Sprong, M., et al. (2009). Proline affects brain function in 22q11DS children with the low activity COMT 158 allele. *Neuropsychopharmacology* 34, 739–746. doi: 10.1038/npp.2008.132

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Psychological Corp.

Wechsler, D. (2003). *WISC-IV Technical and Interpretive Manual*. San Antonio, TX: Psychological Corp.

Zelazo, P. D., Craik, F. I. M., and Booth, L. (2004). Executive function across the life span. *Acta Psychol.* 115, 167–183. doi: 10.1016/j.actpsy.2003.12.005

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 January 2014; accepted: 22 May 2014; published online: 10 June 2014.
Citation: Shapiro HM, Tassone F, Choudhary NS and Simon TJ (2014) The development of cognitive control in children with chromosome 22q11.2 deletion syndrome. Front. Psychol. ***5****:566. doi: 10.3389/fpsyg.2014.00566*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.
Copyright © 2014 Shapiro, Tassone, Choudhary and Simon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 September 2014
doi: 10.3389/fpsyg.2014.00992

Investigating executive functions in children with severe speech and movement disorders using structured tasks

Kristine Stadskleiv[1,2]*, Stephen von Tetzchner[1], Beata Batorowicz[3], Hans van Balkom[4], Annika Dahlgren-Sandberg[5] and Gregor Renner[6]

[1] *Department of Psychology, University of Oslo, Oslo, Norway*
[2] *Section of Paediatric Neuro-habilitation, Department of Clinical Neurosciences for Children, Oslo University Hospital, Oslo, Norway*
[3] *CanChild Centre for Childhood Disability Research, McMaster University, Hamilton, Canada*
[4] *Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, Netherlands*
[5] *Department of Psychology, Gothenburg University, Gothenburg, Sweden*
[6] *Catholic University of Applied Sciences, Freiburg, Germany*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Ayelet Lahat, McMaster University, Canada
Oriane Landry, McMaster University, Canada

****Correspondence:***
Kristine Stadskleiv, Department of Psychology, PO Box 1094 Blindern, 0317 Oslo, Norway;
Section of Peadiatric Neuro-habilitiation, Department of Clinical Neurosciences for Children, Women and Children's Division, Oslo University Hospital, PO Box 4956 Nydalen, 0424 Oslo, Norway
e-mail: kristine.stadskleiv@psykologi.uio.no; kristine.stadskleiv@oslo-universitetssykehus.no

Executive functions are the basis for goal-directed activity and include planning, monitoring, and inhibition, and language seems to play a role in the development of these functions. There is a tradition of studying executive function in both typical and atypical populations, and the present study investigates executive functions in children with severe speech and motor impairments who are communicating using communication aids with graphic symbols, letters, and/or words. There are few neuropsychological studies of children in this group and little is known about their cognitive functioning, including executive functions. It was hypothesized that aided communication would tax executive functions more than speech. Twenty-nine children using communication aids and 27 naturally speaking children participated. Structured tasks resembling everyday activities, where the action goals had to be reached through communication with a partner, were used to get information about executive functions. The children (a) directed the partner to perform actions like building a Lego tower from a model the partner could not see and (b) gave information about an object without naming it to a person who had to guess what object it was. The executive functions of planning, monitoring, and impulse control were coded from the children's on-task behavior. Both groups solved most of the tasks correctly, indicating that aided communicators are able to use language to direct another person to do a complex set of actions. Planning and lack of impulsivity was positively related to task success in both groups. The aided group completed significantly fewer tasks, spent longer time and showed more variation in performance than the comparison group. The aided communicators scored lower on planning and showed more impulsivity than the comparison group, while both groups showed an equal degree of monitoring of the work progress. The results are consistent with the hypothesis that aided language tax executive functions more than speech. The results may also indicate that aided communicators have less experience with these kinds of play activities. The findings broaden the perspective on executive functions and have implications for interventions for motor-impaired children developing aided communication.

Keywords: executive functions, assessment, aided communication, cerebral palsy, severe speech and movement disorder

INTRODUCTION

Executive functions are understood, not as a unitary function but as a psychological construct defined as a set of interrelated high-level cognitive skills that are necessary for purposeful, goal-directed activity (Stuss, 1992; Anderson, 2008; Wiebe et al., 2008; Willoughby and Blair, 2011; Miyake and Friedman, 2012; Benson et al., 2013; Usai et al., 2013). There is a consensus that executive functioning is central in cognitive skills like planning, monitoring results, updating, shifting, and inhibition (Kinsella et al., 2007; Böttcher et al., 2010; Miyake and Friedman, 2012). Planning involves the ability to establish a sequence of sub goals in order to achieve a larger predetermined goal (Hudson and Farran, 2011). Monitoring, or updating (Miyake and Friedman, 2012), involves constant supervision of tasks, with rapid addition and fading of content in working memory. Working memory is the part of the memory system that temporary holds information during mental operations (Eysenck and Keane, 1990; Hitch and Towse, 1995). Inhibition involves overriding of "automatic" behaviors when they are not appropriate (Doebel and Zelazo, 2013). The age at which executive functions emerge is still under debate, but important developments seem to take place from the age of 3 to 4 years (Brocki and Bohlin, 2004; Doebel and Zelazo, 2013).

This has been partly attributed to the emergence of language, which broadens the child's ability to reflect on and reason about the world (Astington and Hughes, 2013). Executive functions are related to daily life skills, academic success, and social functioning (Ganesalingam et al., 2011; Foy and Mann, 2013), and it is therefore important to gain knowledge about how these functions develop in typical and atypical populations. Investigations of atypical development may broaden the understanding of the complex relationships between nature and nurture that drives development (Sameroff, 2010).

According to Luria (1961) and Vygotsky (1986), children's private speech in early childhood helps them in solving difficult tasks. Speech takes on a directing and planning function, and contributes to regulating behavior. Private speech was viewed as a forerunner for inner speech, which is an instrument of the thought process. Later research has confirmed that language plays a role in the development of executive functions, and that the ability to verbalize and name objects supports the performance of executive tasks (Miyake et al., 2004; Landry et al., 2012; Doebel and Zelazo, 2013). However, some children do not develop speech due to severe motor impairments and have to use other means of communication. "Aided communication" is defined as the use of communication aids with graphic symbols (like Pictograms, Picture Communication Symbols, and Blissymbols) or letters and words for face-to-face communication. Graphic symbols and words are used in communication aids, e.g., boards, books, and electronic devices with artificial speech output (von Tetzchner and Martinsen, 2000). The aid vocabulary is organized thematically and hierarchically and the user may have to navigate through several pages to indicate the intended expression(s). Children and adults using aided communication are referred to as "aided communicators" (von Tetzchner and Basil, 2011).

Aided communication is chosen when a child's motor impairment is so severe that the use of speech and manual signs is precluded. Severe motor problems may imply very limited voluntary control over physical actions, including movement of the eyes, the head, the arms, and the legs. Depending on their physical abilities, aided communicators access communication aids either directly or with scanning (Light and Drager, 2007). Direct selection involves any form of pointing, for instance with hand, finger, or eye gaze. Specialized computers with eye gaze technology can detect where on the screen the child is looking (see Higginbotham et al., 2007). Selection with scanning may be independent with the use of switches to control item selection, or with the assistance of a communication partner (partner assisted scanning). If possible, direct selection is the preferred mode of operating a communication aid as this is faster than scanning (Ratcliff, 1994; Light and Drager, 2007). Still, regardless of the access method, it may take aided communicators 1 min or more to name a known object with a single symbol, compared to about 1 s in naturally speaking children (von Tetzchner et al., 2012).

The processes related to constructing utterances with natural speech and graphic symbols are very different. Naturally speaking children produce words with relatively little attention to the articulation itself. The articulation process is automatized and usually requires little monitoring, but problems with speech fluency—when the process from conceptualization to spoken articulation is not running smoothly—has been found to be related to executive problems (Engelhardt et al., 2013). For aided communicators, constructing or "articulating" an utterance involves navigating, using direct selection or scanning, on a communication board or an electronic device with several pages to find and indicate one or more graphic symbols (von Tetzchner and Martinsen, 2000). The aided communicator has to remember the location and find and indicate the graphic symbol(s) expressing the intended meaning. Efficient navigation presupposes knowledge of the structure and organization of the graphic symbols in the communication aid. When constructing aided utterances as fast and precisely as possible, the ability to plan the utterance, monitor the progress, and avoid unnecessary detours through the aid's hierarchical system is important (Oxley, 2003; Murray and Goldbart, 2011; Thistle and Wilkinson, 2013).

The role of speech in regulating behavior has mostly been studied in relation to how children regulate their own actions by using their own speech, and when spoken to Fatzer and Roebers (2012); Landry et al. (2012); Doebel and Zelazo (2013). How children using aided communication first express and then internalize private language expressions is not known, neither how they regulate their own behavior and the behavior of others through language, as performing complex actions to reach a goal might be unavailable for them due to their physical impairments. However, language can also be used to regulate the behavior of other people and a child may use language to make others perform actions to reach a particular goal. In such situations, the child's effective use of language implies the use of executive functions. Exploring how young naturally speaking and aided communicators make plans, monitor progress and avoid impulsive errors while using language to direct the actions of others to obtain a goal may therefore give insight into the relationship between language and executive functions.

Using a communication aid requires conscious navigation and deliberate monitoring and the motor impairments of many aided communicators tend to prevent automation of the selection process, thereby placing a constant demand on working memory (Oxley, 2003) and other aspects of executive functioning. Executive functions are generally involved in the construction of aided utterances but the demand on them may vary with communication mode. When an aided communicator is using graphic symbols in a communication book to construct an utterance, the demands on working memory, planning and monitoring will be high, and the avoidance of impulsive errors may be difficult. Also utterances constructed with letters may take a long time to produce but the need for organization and planning might be less with spelling than with graphic symbols because the number of letters is limited and the letters are usually visible all the time.

However, it is usually not only expressive language construction that is affected in aided communicators. Severe physical impairments may make aided communicators unable to reach a physical goal with their own motor acts. Their only means of acting on the physical world may be through instructing other people to perform the actions, that is, by using language for action (Batorowicz et al., 2013). Language may thus have a more decisive role in play and other activities for children with severe motor impairments than for children without such difficulties.

Motor impairment may influence a child's experiences in several ways. Studies show that hands-on experience also contributes to children's regulation of their behavior. One study found that learning to say name shapes like "rhomboid" and "triangle" was not sufficient to make children differentiate between them; they needed the additional information that was gained from touching the shapes (Luria, 1961). Children's participation in social interaction is also believed to influence the development of higher mental functions (Vygotsky, 1986). Children with severe motor and speech impairments have both less experience with handling physical objects than their peers and fewer social experiences (Caillies et al., 2012). In conversations involving aided communicators, the communication partner often takes the initiative, decides the topic and asks questions that only require yes and no answers (von Tetzchner and Grove, 2003; Falkman et al., 2005; Ferm et al., 2005; Clarke et al., 2012). The developmental consequences of these experiences are not known, but from a neuroconstructivist perspective (Mareschal, 2011; Böttcher, 2012) it seems likely that the developmental trajectories of cognition, language, and social functions in the children are negatively affected.

For aided communicators, executive functioning thus seems important both in the construction of utterances and when striving to reach action goals through the use of instructional language. However, there is very little research on executive functioning in this group, and the consequences severe speech and movement disorders may have for the development of executive functions is not known. Indeed, cognitive functioning in general has been little studied in this group, apart from studies looking at the prevalence of intellectual disability (e.g., Andersen et al., 2008; Beckung et al., 2008; Sigurdardottir et al., 2008). Studies investigating different cognitive functions in children with severe speech and movement disorders are therefore needed.

Studies of less motor disabled and mainly speaking children with cerebral palsy (CP) have found that tasks that make demands on executive functioning may be challenging (White and Christ, 2005; Jenks et al., 2007; Böttcher et al., 2010; Pirila et al., 2011; Whittingham et al., 2014) but with some variation with regard to which functions that are affected. Working memory has been found to be reduced, but not inhibitory control (Caillies et al., 2012). However, these studies have rarely included the third of the CP population who are severely motor impaired and in need of aided communication (Andersen et al., 2010). The few studies of severely speech and motor impaired children have focused on attention and working memory, and have found that their attention was reduced compared to peers matched for mental age (Dahlgren et al., 2010), that visual and spatial working memory but not phonological memory was affected (Larsson and Dahlgren Sandberg, 2008), and that working memory capacity increased less than expected from 6 to 12 years (Dahlgren Sandberg, 2006). Several authors have discussed the role of working memory and executive functioning in aided communicators (Light and Lindsay, 1991; Ratcliff, 1994; Oxley and Norris, 2000; Oxley, 2003; Murray and Goldbart, 2011) but to our knowledge there are still no empirical studies of other executive functions than working memory. Indeed, the effectiveness of the aided communication process itself has hardly been investigated (Novak et al., 2013) and little is known about how aided communicators construct utterances that are more complex than when making binary choices such as choosing between milk and juice or between listening to music and watching a video (Murray and Goldbart, 2009). There are studies of adults with aphasia showing that executive functions influence strategy use in communication tasks (Purdy and Koch, 2006), including the learning of graphic symbols (Nicholas et al., 2011), but studies of adults with aphasia may have limited relevance for understanding the development of executive functioning in children with motor impairment.

Measuring executive functions is usually done by either neuropsychological assessment (Lezak, 2004; Strauss et al., 2006), questionnaires (Egeland and Fallmyr, 2010; McCoy et al., 2011) or behavioral tasks (Bechara et al., 1994; Carlson, 2005), or a combination of these. Assessing executive functions in children with severe speech and movement disorders with neuropsychological instruments meets with some challenges, as tests that require the ability to draw, manipulate objects or give a rapid verbal or motor responses cannot be used when the children are unable to perform such actions (Schiørbeck and Stadskleiv, 2008). The tests would therefore need to be adapted, for instance by altering response modality (Alant and Casey, 2005). No validated versions of adapted neuropsychological measures of executive functions exist as of today. Questionnaires such as Behavior Rating Inventory of Executive Functions (Gioia et al., 2000) presupposes that the child can move and talk without restraint. It is therefore necessary to find ways of assessing executive functions that are suitable for children with severe speech and movement impairments (Clarke et al., 2012).

There is a long tradition of using behavioral tasks for exploring aspects of executive functions (see Carlson, 2005), and it has been argued that such tasks may reflect real-life functioning better than test items (Bechara et al., 1994). When using tasks to investigate executive functions in aided communicators, it is important that the tasks draw on the child's best skill, which is language.

The present study thus investigates executive functioning in children with severe speech and movement disorders by using two tasks that resemble everyday activities that require executive functions. In the first task, the child instructs a partner to perform a complex set of actions, such as building a tower of blocks, instead of the child himself executing the actions. In the second, the child is instructed to describe, but not name, objects. Together, these tasks require planning, monitoring of the progress of the task, and avoiding impulsive errors. The executive functions that are required to plan how to instruct the partner so that the building of the tower goes as efficiently as possible, the monitoring of the construction process that is required to correct any misunderstandings and the need to inhibit the impulse to name the objects, thereby breaking the rules of the task, are comparable to the executive functions that are tapped using tests like tower tests and the Stroop test (see Delis et al., 2001). With this approach, executive functions are investigated through language instead of the performance of physical actions. This also reduces the problems encountered when trying to adapt standard neuropsychological tests to children with severe speech and movement impairment.

The performance of the aided communicators is compared to that of typically developing and naturally speaking peers. It was

hypothesized that the aided communicators would have more problems and errors than the comparison group, as a consequence of the aided group having a double load on executive functions when performing everyday tasks through others, that is, executive functions both in planning, organizing, and monitoring the tasks itself, and simultaneously in constructing aided utterances. It was further hypothesized that the aided communicators using graphic symbols would have more difficulties than the children using spelling as their communication mode, as the demands on planning and organization is assumed to be higher when navigating through a communication aid than when spelling.

MATERIALS AND METHODS

This study is part of the international project "Becoming an aided communicator (BAC). Aided language skills in children aged 5–15 years—a multi-site and cross-cultural investigation," which involves children from 16 countries (von Tetzchner et al., 2012). The present study reports on the performance of aided and naturally speaking communicators in Norway, Canada, the Netherlands, Sweden, and Germany on tasks that they solve in dialog with a communication partner.

PARTICIPANTS

The children using aided communication were recruited with the help of professionals in the specialized healthcare system and special education system in each of the regions. A search was made for all the high-functioning aided communicators who met the following criteria: (a) were between ages 5 and 15; (b) had used communication aids for a minimum of 1 year; (c) had normal hearing and vision (with corrective technology); (d) were not considered intellectually impaired by their teacher; (e) did not have a diagnosis in the autism spectrum; (f) had speech comprehension considered adequate or near adequate for their age; and (g) speech production was absent or very difficult to understand.

The comparison children were recruited from the class of the aided communicator or from the closest school in the same type of neighborhood (e.g., rural or urban) if the aided communicator went to a special school. The comparison child had the same gender and was the student closest in date of birth to the aided communicator. All children in the comparison group were using speech as communication mode, were healthy, had normal vision and hearing, had no known learning disabilities and attended mainstream schools.

Twenty-nine children using aided communication and 27 typically developing children participated. 20 of the children were Norwegians, 15 Canadian, 12 from the Netherlands, five Swedish, and four from Germany. The age of the children spanned from 6;7 to 15;11 years. Eleven children in each of the groups were boys (see **Table 1**). There was no significant difference between the aided and the comparison group on age or gender. Twenty-seven of the aided communicators had a diagnosis of CP. For the remaining two children the diagnosis was unknown.

Language comprehension was assessed with the Test for the Reception of Grammar, second edition (TROG-2) using national norms (Bishop, 2003) (see **Table 1**). The test places small demands on motor skills; the child indicates which of four pictures that corresponds to the sentence spoken. If the child is unable to point with the hand, a partner used assisted scanning. This implies that the four pictures were pointed at in a systematic manner and the child indicated "yes" or "no" for each picture (Schiørbeck and Stadskleiv, 2008). The mean scores were below the age mean, but within two standard deviations of the mean, for both groups. The difference between the groups was significant.

Naming and expressive language speed was compared using *BAC Naming*. The task of the child is to name drawings of 20 objects and animals, which are shown one at a time. The aided communicators correctly named on average 14.2 (71%) of the drawings and the comparison group 19.6 (98%), a significant difference. The aided group used significantly longer time giving correct names than the comparison group. There is a significant and positive correlation ($r = 0.60$) between scores on TROG-2 and number of correct answers on BAC Naming.

A description of the aided communicators is given in **Table 2**. Gross and fine motor functions were classified according to the Gross Motor Classification System (GMFCS) (Palisano et al., 1997) and Manual Ability Classification System (MACS) (Eliasson et al., 2006). They both have five levels with level I indicating best possible functioning and level V severe difficulties. On level I of GMFCS, the child walks without assistance, while on level V the child cannot sit or stand independently. On level I of MACS, the child handles objects easily and successfully, while on level V the child is unable to handle objects. The median scores of the aided communicators were level V on both GMFCS and MACS.

Table 1 | Characteristics of the participants.

	Aided	Comparison	*t*	*p*
	Mean (*SD*)	Mean (*SD*)		
Age, months ($N = 29, 27$)	136.1 (33.7)	137.1 (31.5)	−0.114	0.910
Gender, boys % ($N = 29, 27$)	37.9 (49.4)	40.7 (50.1)	−0.211	0.834
TROG-2, z-scores ($N = 28, 14$)	−1.50 (1.22)	−0.29 (0.99)	−3.434	0.002*
BAC Naming, correct out of 20 items ($N = 23, 14$)	14.2 (5.5)	19.6 (0.6)	−4.641	0.000*
BAC Naming, mean time in sec ($N = 20, 14$)	42.8 (32.0)	1.8 (0.7)	5.568	0.000*

t, t-test (independent samples t-test, equal variances not assumed).

*$p < 0.05$.

Table 2 | Characteristics of the aided communicators.

Variable	Specification	Aided group		*t*	*p*
		Symbols users	Spellers		
Gender, boys % (*N* = 15, 14)	Mean (*SD*)	26.7 (45.8)	50.0 (51.9)	−1.281	0.212
Age, in months (*N* = 15, 14)	Mean (*SD*)	120.5 (37.0)	152.9 (19.6)	−2.972	0.007*
GMFCS level (*N* = 15, 14)	Level I (n)	1	0		
	Level II (n)	0	0		
	Level III (n)	2	0		
	Level IV (n)	5	4		
	Level V (n)	7	10		
MACS level (*N* = 15, 14)	Level I (n)	1	0		
	Level II (n)	1	0		
	Level III (n)	2	0		
	Level IV (n)	4	4		
	Level V (n)	7	10		
Viking speech scale level (*N* = 15, 14)	Level I (n)	0	0		
	Level II (n)	0	0		
	Level III (n)	1	0		
	Level IV (n)	14	14		
CFCS level (*N* = 11, 12)	Level I (n)	0	0		
	Level II (n)	3	10		
	Level III (n)	6	2		
	Level IV (n)	2	0		
	Level V (n)	0	0		
Communication access (*N* = 10, 10)	Direct (hand) (n)	7	1		
	Direct (gaze) (n)	2	2		
	Scanning, switches (n)	1	7		
TROG-2, z-scores (*N* = 15, 13)	Mean (SD)	−2.09 (1.04)	−0.82 (1.08)	−3.154	0.004*
BAC Naming, correct of 20 items (*N* = 11, 12)	Mean (SD)	9.7 (4.3)	18.3 (2.4)	−5.749	0.000*
BAC Naming, time in sec (*N* = 10,10)	Mean (SD)	49.3 (43.6)	36.4 (17.5)	0.873	0.400

t, t-test (independent samples t-test, equal variances not assumed).

*$p < 0.05$.

The Viking Speech Scale (Pennington et al., 2010) has four levels, with level I indicating normal function and level IV being used for children with no functional speech. The median scores of the aided communicators was level IV. The Communication Functioning Classification system (CFCS) (Hidecker et al., 2011) has five levels with level I indicating best functioning and level V the most affected. On level I the child can communicate efficiently and without reduced speed with both known and unknown partners, on level II the child can communicate efficiently with both familiar and unfamiliar partners, but the speed of communication is reduced compared to peers, on level III the child can communicate effectively with known partners, but not with unknown, on level IV the child can communicate somewhat efficiently with known partners, while children on level V have difficulties with being understood even by familiar partners. The aided communicators had scores between II and IV, with a median score of II for the spellers and III for the symbol users.

Fourteen of the 29 children used spelling, either alone or in combination with graphic symbols, while 15 used only graphic symbols. Most of the symbol users accessed their communication devices directly by pointing with hand or eye gaze. Of the spellers, scanning with two switches (i.e., step scanning, with one switch used to progress between items and a second to select the item) was most common. There were no significant differences between children using symbols and children using spelling concerning gender and speed of naming objects, but the spellers were on average more than 2 years older, showed better comprehension of language and named significantly more objects correctly than the non-spellers.

Parents, peers, and teachers of the aided and the comparison group were asked to participate as communication partners in the study. The children in the aided group and the comparison group were asked to nominate a peer with whom they wanted to do the tasks. The peers were friends whom they knew well, and the aided communicators and the peers had experience in communicating together. Some of the children in the aided group were unable to nominate a friend, and a sibling was asked instead. The parents of the all the participating children (aided group, comparison group and peers) gave consent to their child's participation.

TASKS

The two tasks *BAC Construction* and *BAC Description without naming* were used to obtain measures of planning, monitoring, and impulsivity. The performance on the tasks was videotaped and the dialog was transcribed in accordance with existing standards for such transcriptions (see von Tetzchner and Basil, 2011). A coding system was developed based on a detailed analysis of the videos of four children, each interacting with three different partners. Inter-rater reliability was established by two independent raters who watched the videos and scored all tasks until full agreement was reached.

BAC Construction

In BAC Construction the child has a physical model in a box, placed so that the model is not visible to the partner. The child instructs the partner to construct the same figure. There are two training items (loading a lorry) and eight task items (dressing a doll, making a necklace of beads, building a tower of Lego blocks, and laying a pattern of domino pieces). The partner has many more clothes, beads, Lego blocks, and domino pieces than are needed for construction. To reduce the memory load of the child, the model was visible to the child throughout the task.

Table 3 shows the measures used in the analysis: correctly solved items, the time it took to solve the items, the child's planning ability, as well as the child's monitoring of the construction process (see also Batorowicz et al., 2013, 2014).

Items solved correctly are items where the partner constructed an exact copy of the child's figure. Almost similar models, like Lego towers with one block in the wrong color, were scored as failed.

Planning is defined as the type of strategy that could be observed when the child provided instructions to the partner. The quality of the child's planning was classified on a scale from 0 to 3. A score of 0 means that it was difficult to decide if the child had a plan, a score of 1 means that the child did not seem to have a clear plan (e.g., if the child started the Lego item by describing a block positioned in the middle of the tower), a score of 2 means that the child initially did not seem to have a clear plan, but that a plan seemed to emerge during the item performance (e.g., started by having the partner put on shoes before trousers on the doll, but then progressed without problems from there on), and a score of 3 means that the child seemed to have a clear plan throughout the item. A score for the observed planning was given for each of the

Table 3 | Task measures.

Task	Measures	Categories		Scores
BAC Construction	Items correctly solved	Items constructed correctly		0–100%
	Time	Average time use per item (sec)		
	Planning	Difficult to decide if the child has a plan		0
		No evidence of a plan		1
		Plan evolved during solution of task		2
		Clear plan from the beginning		3
	Object	Proportion of necessary objects named	Less than necessary number of objects mentioned	<1.00
			Necessary number of objects mentioned	1.00
			More than necessary number of objects mentioned	>1.00
	Attributes	Proportion of necessary attributes named	Less than necessary number of attributes mentioned	<1.00
			Necessary number of attributes mentioned	1.00
			More than necessary number of attributes mentioned	>1.00
	Specificity	Wrong description		1
		Too low specificity (only correct object or attributes)		2
		Perfect specificity (correct object and attributes)		3
		Too high specificity (correct object, one attribute too many)		4
		Much too high specificity (correct object, too many attributes)		5
	Monitoring	Specificity provided at the start of the item subtracted from specificity provided after watching the partner constructing		0.0–5.0
BAC Description	Items correctly solved	Items named correctly by partner		0–100%
	Impulsivity	Items the child names, contrary to instruction		0–100%

eight items administered, and the child's average planning score for the whole task was based on this.

A five-point *specificity scale* was used to measure *monitoring*, based on the preciseness of the child's utterance. The measure makes it possible to look not only on the quantity of information provided, but also on the quality of it. A precise description should include both a description of the *object* (such as the type of clothes needed to dress the doll) and the *attributes* of the object (e.g., color of the pants). The number of objects and attributes the child mentioned was compared to the number that was necessary for a precise description. A specificity score of 1 would indicate very low specificity (wrong description), a score of 2 a little too low specificity (only correct object or only correct attributes), a score of 3 a perfect specificity (the objects and attributes needed), a score of 4 a little too high specificity (one attribute too many), and a score of 5 a much too high level of specificity (more than one attribute too many). *Initial specificity* is the score of the content provided by the child before the partner has started constructing, while the *final specificity* is the total content provided by the child while watching the partner constructing the item. If the child did not add any information during the task, the final specificity would be equal the initial. If the child saw the need to add information while watching the partner constructing, the final specificity would be higher than the initial. *Monitoring* is the difference between initial and final specificity. It is a continuous variable ranging from 0 (no new information added) to a maximum of 5 (the maximum specificity score); If the child adds any task-relevant information while observing the partner constructing the item, the monitoring score will be a positive number larger than 0 and smaller than 5.

BAC Description without naming

In "BAC Description without naming" (henceforth abbreviated as BAC Description), the child was presented with 12 different drawings of an object and instructed to describe but not name, the objects in such a manner that the communication partner could name the objects. Three of these 12 drawings were training tasks, so that a total of nine drawings are included in the analysis. The drawings were in a box and not visible to the partner who was seated opposite the child. The partner was allowed to make as many guesses as were necessary to name each item, but not to ask any leading questions (like "What is it used for?"). The child could continue to describe the object also after the partner had started to guess. To reduce the memory load of the child, the drawing was visible to the child throughout the task. The three training items and nine task items consisted of common objects like a chair, a book, and an apple. From this task measures of correctly solved items and impulsivity were obtained (see **Table 3**).

Correctly solved items are items where the partner names the object. Items that were almost solved, for example saying "orange" when there was a picture of an apple, were scored as failed, as were items where the partner answered wrong or was not able to make a guess.

Impulsivity is when the child names the object. This is a violation of the task instruction as the child was instructed not name the object.

ETHICS

The ethical approval for the study was obtained by each participating site following the national procedures for ethical approval.

STATISTICS

Data was coded in IBM SPSS Statistics, version 20. Independent sample *t*-tests were used for comparisons between the aided group and comparison group, and between aided communicators using symbols and letters. Spearman's Rho rank order correlations were used to investigate the relationship between variables.

RESULTS

There was a significant correlation between age and percentage of correctly solved items on the BAC Construction task, but not on the BAC Description task (see **Table 4**) when looking at both the aided group and the comparison group together. Verbal comprehension, as measured by results on TROG-2 was not significantly related to percentage of solved items on the BAC Description task, but was related success on the BAC Construction task. Expressive verbal abilities, as measured by number of correct items on the BAC Naming task, were significantly and positively correlated with percentage of solved items both on the BAC Construction task and on the BAC Description task.

There were differences between the aided group and the comparison group with regards to task success. On BAC Construction, the aided group solved 64.7% of the items and the comparison group 92.6%, a significant difference (see **Table 5**). The aided communicators took almost five times as long as the naturally speaking children to complete the items on this task. On the BAC Description task, the aided group solved 65.1% and the comparison group 96.7%, a significant difference.

PLANNING, MONITORING, AND IMPULSIVITY IN AIDED COMMUNICATORS AND COMPARISON GROUP

A significant group difference in the planning score (see **Table 5**) indicates that more children in the aided group than in the comparison either did not have a plan from the start or had to develop a plan during the item. The aided group described significantly less objects and attributes than the comparison group. In both groups there was a considerable variation in number of attributes mentioned, with means ranging from 0.28 to 2.76 in the aided group and from 0.78 to 2.79 in the comparison group. (A proportion larger than 1.00 indicates that more attributes

Table 4 | Correlation between items correctly solved and age, verbal comprehension and expressive verbal abilities.

	BAC Construction items correctly solved	BAC Description items correctly solved
Age	0.37**	0.27
Verbal comprehension	0.45**	0.24
Expressive verbal abilities	0.72**	0.46**

*$p < 0.05$, Spearman's Rho rank order correlations (two-tailed).

**$p < 0.05$, Spearman's Rho rank order correlations (two-tailed).

than necessary were described.) In the aided group, the initial and the final specificity were significantly below the initial and final specificity in the comparison group. The increase in specificity from initial to final was equal in both groups, implying that there was no group difference with regards to monitoring. On the BAC Description task, there was significantly more evidence of impulsivity in the aided group than in the comparison group.

PLANNING, MONITORING, AND IMPULSIVITY IN RELATION TO COMMUNICATION MODE

On the BAC Construction and the BAC Description tasks the percentages of solved items were significantly lower for aided communicators using graphic symbols than for those using spelling (see **Table 6**). There was a significant difference between the two groups in planning and specificity of the utterances, but not on monitoring or degree of impulsivity.

EXECUTIVE FUNCTIONING AND TASK SUCCESS

Aspects of executive functioning correlated significantly with task success when the results for both groups were combined (see **Table 7**). In the combined group (aided communicators and comparison group) there was a significant positive correlation between planning and number of correct items on the tasks and a negative correlation between impulsivity and percentage of correct items on both tasks. Monitoring and performance on the BAC Description task was significantly and negatively associated in the aided group and in the combined group.

Table 5 | Task performance of aided group and comparison group.

Tasks and variables	Aided	Comparison	*t*	*p*
	Mean (*SD*)	Mean (*SD*)		
BAC CONSTRUCTION				
Items correctly solved (%) ($N = 28, 27$)	64.7 (34.1)	92.6 (10.1)	−4.147	0.000*
Time (sec) ($N = 26,27$)	419.6 (251.7)	88.2 (51.7)	6.808	0.000*
Planning ($N = 27, 27$)	2.5 (0.5)	2.8 (0.2)	−3.353	0.002*
Objects ($N = 27, 27$)	0.58 (0.29)	0.90 (0.10)	−5.384	0.000*
Attributes ($N = 27, 27$)	1.01 (0.51)	1.67 (0.55)	−4.665	0.000*
Specificity of utterance (initial) ($N = 27, 27$)	2.2 (0.6)	3.2 (0.5)	−6.575	0.000*
Specificity of utterance (final) ($N = 27, 27$)	2.6 (0.7)	3.5 (0.5)	−6.029	0.000*
Monitoring ($N = 28, 27$)	0.4 (0.2)	0.3 (0.3)	0.693	0.491
BAC DESCRIPTION				
Items correctly solved (%) ($N = 27, 22$)	65.1 (28.7)	96.7 (7.4)	−5.517	0.000*
Impulsivity (%) ($N = 27, 25$)	14.6 (23.9)	1.6 (4.4)	2.793	0.009*

t, t-test (independent samples t-test, equal variances not assumed).
*$p < 0.05$.

Table 6 | Task performance of aided communicators using symbols and spelling.

Tasks and variables	Symbols	Spelling	*t*	*p*
	Mean (*SD*)	Mean (*SD*)		
BAC CONSTRUCTION				
Items correctly solved (%) ($N = 15, 13$)	43.9 (33.2)	88.6 (12.9)	−4.810	0.000*
Time (sec) ($N = 15, 13$)	353.7 (233.0)	495.6 (259.8)	−1.511	0.144
Planning ($N = 14, 13$)	2.2 (0.5)	2.8 (0.4)	−3.161	0.004*
Objects ($N = 14, 13$)	0.44 (0.29)	0.74 (0.20)	−3.131	0.005*
Attributes ($N = 14, 13$)	0.92 (0.44)	1.10 (0.57)	−0.950	0.352
Specificity (initial) ($N = 14, 13$)	1.9 (0.6)	2.6 (0.5)	−3.637	0.001*
Specificity (final) ($N = 14, 13$)	2.2 (0.7)	2.9 (0.5)	−3.022	0.006*
Monitoring ($N = 14, 13$)	0.4 (0.2)	0.3 (0.2)	0.733	0.471
BAC DESCRIPTION				
Items correctly solved (%) ($N = 13, 14$)	52.5 (26.1)	76.8 (26.6)	−2.402	0.024*
Time (sec) ($N = 4, 8$)	234.2 (177.8)	247.9 (139.6)	−0.134	0.898
Impulsivity (%) ($N = 13, 14$)	20.1 (20.0)	9.6 (26.7)	1.162	0.257

t, t-test (independent samples t-test, equal variances assumed).
*$p < 0.05$.

Table 7 | Correlation between executive functions and items correctly solved on BAC Construction and BAC Description.

	Aided group		Comparison group		Groups combined	
	Items correctly solved		Items correctly solved		Items correctly solved	
	BAC construction	BAC description	BAC construction	BAC description	BAC construction	BAC description
Planning	0.64**	0.37	0.26	0.19	0.54**	0.42**
Monitoring	0.16	−0.46*	0.08	−0.34	0.05	−0.38**
Impulsivity	−0.34	−0.63	−0.42	−0.88**	−0.39**	−0.69**

*$p < 0.05$, Spearman's Rho rank order correlations (two-tailed).
**$p < 0.01$, Spearman's Rho rank order correlations (two-tailed).

Table 8 | Correlation between executive functions and language comprehension and expressive language skills.

	Aided group		Comparison group		Groups combined	
	Language comprehension	Expressive verbal abilities	Language comprehension	Expressive verbal abilities	Language comprehension	Expressive verbal abilities
Planning	0.40*	0.60**	−0.30	0.38	0.38*	0.54**
Monitoring	0.09	0.03	−0.33	0.07	0.09	0.05
Impulsivity	−0.13	−0.26	0.22	0.21	−0.21	−0.33

*$p < 0.05$, Spearman's Rho rank order correlations (two-tailed).
**$p < 0.01$, Spearman's Rho rank order correlations (two-tailed).

EXECUTIVE FUNCTIONING AND VERBAL ABILITIES

Planning was positively related to verbal comprehension and expressive verbal abilities in the aided group, but not in the comparison group (see **Table 8**). Monitoring and impulsivity were not related to verbal comprehension or expressive verbal abilities, neither in the aided group nor the comparison group alone, or in the groups combined.

DISCUSSION

Both the aided group and the comparison group completed most of the items in the BAC Description and BAC Construction tasks correctly. In spite of their severe motor disabilities and lack of speech, the aided communicators were able to plan, execute and monitor instructions to make partners perform the physical acts needed to construct something the partners could not see. They were also able to describe objects in such a way that the partners, who could not see the object, were able to name them. This demonstrates how language use can compensate for the lack of motor skills that are required to act directly on the physical world and the achievements of the motor-impaired aided communicators. However, the considerable time and effort the aided communicators needed to complete the tasks compared to their typically developing peers imply a continuous demand on executive functions.

The aided group was hypothesized to have more difficulties with the tasks because planning both the language construction and the complex set of actions for the partner to perform represents a double demand on executive functions. The naturally speaking children only had to plan the actions as articulation of speech was automatized and required little cognitive effort. It was also hypothesized that the aided communicators using spelling would perform better than the aided communicators using graphic symbols, because finding and selecting symbols in a communication book or electronic communication aid place larger demands on executive functions than spelling. Both of these hypotheses were supported: the aided communicators solved significantly fewer tasks than the comparison group and the graphic symbol users solved significantly fewer tasks than the spellers. The results indicate that although language use can compensate for being unable to perform a complex set of actions to reach a physical goal (language for action), the process of aided language construction, and especially when involving the use of graphic symbols, was taxing the children's overall executive capacity.

The ability to make a clear plan was positively related to outcome on both the tasks for the groups combined, and on the BAC Construction task in the aided group alone. This may reflect that for aided communicators the creation of an initial plan and how they communicate this from the start plays a greater role for task success than in naturally speaking children who can correct mistakes more easily while monitoring the construction process. The results, specifically the scores on number of objects and attributes mentioned and the specificity scores show that on average the comparison group provided more information than the aided group. This may reflect the ease of articulation of speech compared to constructing aided utterances. As the specificity scale gives an evaluation of the quality of the information provided, not only of the amount, the results show that the aided communicators provided somewhat imprecise information at the start of the task and then added necessary information, while the comparison group on average provided precise enough information at the start but still added more details while observing the partner's task construction.

Monitoring was defined as the child providing more information to the partner after the initial description and as a result of watching the construction process, and both groups provided the same amount of extra information. However, in the aided group, monitoring was negatively related to outcome on the BAC description task. This negative relationship might reflect that for the children who needed to supply the most extra information on the construction task, the task of describing an object without any support from a communication partner was extra challenging.

Monitoring might also be viewed as an intrinsic part of the aided communication process. However, aided communicators do not have the full responsibility for this because aided language production tend to be co-constructive; that is, the naturally speaking communication partner makes interpretations and guesses during message construction which are confirmed or rejected by the aided communicator (Collins, 1996; von Tetzchner and Grove, 2003). When aided communicators construct a graphic utterance with symbols the communication partner usually interprets the meaning of each symbol, formulates the utterance in speech and seeks acknowledgement of the spoken formulation from the aided communicator. The partners thus function as interpreters and translators. This is true also when the utterance is produced with artificial speech, unless the graphic symbol or symbol sequence produces a pre-made sentence. Moreover, although it is the aided communicator who has to produce the graphic utterance, the communication partners often take a leading role and dominate the co-construction even when the message is about an event that is unknown to them (von Tetzchner and Martinsen, 1996; Collins and Marková, 1999; Batorowicz et al., 2014). Monitoring is thus a core element of the aided language process and this might explain why monitoring was less affected by aided language experience. This also implies that the emergence of language and the language construction process is quite different in aided and naturally speaking communicators and that the many proposed mechanisms in typical language construction and development (see Gerken, 2005) may not apply to the same extent in the construction and development of utterances with communication aids. Utterances with graphic symbols are produced, but may also be processed and represented mentally, in a different manner from speech.

Impulsivity was negatively related to outcome in both groups. There was more evidence of impulsivity in the aided group than in the comparison group, but most of the children in both groups had no problems with inhibiting the impulse to name an object they were instructed not to name. This is in line with findings from other studies, indicating that inhibition is not affected in children with CP (Caillies et al., 2012). There was, however, some variation within the aided group, where the graphic symbol users made somewhat more impulsive errors than the spellers, that is, they named more objects instead of describing them. One reason might be that communication aids with graphic symbols contain a fixed vocabulary that can be used for constructing utterances. A limited number of graphic symbols imply that each symbol may have to be used to construct a broader set of meanings than the corresponding word for speaking children. One result of this may be a co-construction process where aided communicators provide one or two key words and then rely on their communication partner to complete the message (von Tetzchner and Martinsen, 2000; Brekke and von Tetzchner, 2003). Children who are able to spell are less restricted in producing their utterances than graphic symbol users. However, the results also show that with clear instructions and training most of the symbol users were able to provide a richer description than just naming.

Monitoring and impulsivity were not related to language comprehension or expressive language skills, but to overall success on the tasks. Planning was also related to success on the tasks, as well as to language comprehension and expressive language skills. These findings indicate that not only comprehension, but also the quality of the verbal output of the child, correlates with the regulation of behavior. Related findings have been reported in other studies, where expressive language was found to play a role when children were asked to regulate their own behavior (Fatzer and Roebers, 2012; Landry et al., 2012; Doebel and Zelazo, 2013) and verbal fluency in the development of executive functions (Brocki and Bohlin, 2004).

In addition to the demands on planning, the children's ability to use language to regulate another person's behavior may have been influenced by their earlier experiences. Compared to typically developing children, children with motor impairments are likely to have less experience with active involvement in ordinary construction play like dressing a doll or building a tower of blocks or other construction activities (Caillies et al., 2012) and hence have fewer experiences to build on when trying to find ways to do the tasks in an efficient manner. Instructing others may compensate for the child's lack of motor skills but descriptions of child-adult interactions where the child instructs the adult to do something that is unknown to the adult are rare (see von Tetzchner and Martinsen, 1996). Aided communicators are therefore likely to have limited experience with giving others detailed instructions to construct something. The comparison group probably had more experience with construction play. They may also have relatively little experience with this form of instructing actions but explaining to others how to do things is not uncommon among children (for instance in play activities). This finding is in line with the earlier cited finding that verbal skills are not sufficient to solve a task, but that hands-on experience is also needed (Luria, 1961).

The results show that there is more variation in the aided group than in the comparison group. This may indicate that the aided group was more heterogeneous than intended when searching for participants. All the aided communicators were judged by their teachers not to be intellectually impaired and the results on TROG-2 supported this evaluation as the group mean was within two standard deviations from the age mean. It is, however, possible that some of the aided communicators had specific difficulties in some areas which were not discovered by the teachers. There seemed to be some correspondence between communication mode and language level and age. The graphic symbol users were younger than the spellers, which was expected as the expressive language development of aided communicators typically starts with graphic symbols and progress to spelling (although many continue to have problems with reading and writing). They also scored lower on verbal comprehension and naming. However, a thorough assessment of aided communicators

is recommended, so that specific cognitive challenges can be detected and taken into consideration in educational planning.

SUGGESTIONS FOR INTERVENTIONS

Through development there is a complex interplay between neurological impairment and plasticity, cognitive development, and developmentally affording experiences (Böttcher, 2010). Using language for action may compensate for motor impaired children's limited physical and social experiences, but this implies good abilities for planning and organization. One recommendation that emerges from the findings in this study is to support the development of executive functions by giving young aided communicators more opportunities for engaging in construction play and other construction activities. This may support not only executive functions, but also lead to greater autonomy and social participation (Batorowicz et al., 2014).

Furthermore, greater focus may be given to structural stability in aid content to promote increasing automation of language construction in children using aided language. Spelling might reduce the executive demands inherent in the construction process but many nonspeaking children have significant difficulties and delays related to literacy acquisition independent of general cognitive function (Smith, 2005; Larsson and Dahlgren Sandberg, 2008), so this might not be an available option for all aided communicators. It is also notable that the ability to spell did not lead to faster solutions, implying that this form of communication mode is still taxing on the child's cognitive capacity.

LIMITATIONS OF THE PRESENT STUDY

The present study includes a small group of aided communicators and a matched group of typically developing peers. It cannot be ruled out at that the small sample size may have caused a bias in one way or the other, and the findings of the present study will need to be confirmed in future studies.

A search for aided communicators filling the inclusion criteria was made in all the regions included in the study. Although the aim was not to have a complete geographical sample, the experience was that finding the aided comparison children proved harder than anticipated at the start of the study. This might indicate that the group presented in this study is representative for the high functioning aided communicators. For instance, for the Norwegian sample children from the 11 southernmost counties of Norway, where approximately 60% of the country's population live, were included. Previous studies in Norway have indicated that 0.023% of children use graphic communication and that a quarter of these children can be classified as belonging to the expressive group (von Tetzchner, 1997). With a population of 5 100 000, where 12.1% is children between 6 and 15 years of age (Statistics Norway, 2014[1]), it can be estimated that the expressive group in the region included in the study totals approximately 21–22 children. So even though the sample size is small, it probably represents a fair proportion of the eligible children in the geographical area covered.

[1]http://www.ssb.no/befolkning/statistikker/folkemengde/aar/2014-02-20#content

The measures that were used in this study give insight into the use of executive functions in daily life as the tasks chosen resemble everyday activities that children are likely to encounter, such as building a tower of blocks, making a pattern with beads, matching amounts and telling people about something they have observed but the other person has not. The ecological validity of the tasks is therefore assumed to be high. As we have not employed other methods for investigating executive functions, the study cannot give information about how these measures compare to other measures of executive functions, such as neuropsychological tests.

SUGGESTIONS FOR FURTHER STUDIES

The study has shown that structured tasks resembling everyday activities can give important information about executive functions in a group of children where this information is particularly important and for whom traditional tests and questionnaires cannot be used without adaptations. The tasks required very few skills besides the ability to produce an utterance (in any manner possible and without any time limit) and utilized the area where children with severe motor impairments function best. Further studies should look at the kind of interventions that are provided to this group of children and how these may influence development and learning in aided communicators. In spite of the potential importance of being able to use language to guide another person to perform a specific set of actions, this way of using aided communication is hardly described. An important research issue might be to develop interventions aimed at providing opportunities for creative language construction and active exploration of the environment, and investigate their influence on executive functioning.

The results for the comparison children, who mastered nearly all the items, point to the possibility that the items may have been too easy to give information about executive functions in a group of typically developing 5–15 years olds. For instance, in the Lego task they only needed to describe a model of a tower and not a more challenging three-dimensional model. Future studies should take this into consideration.

Future studies may also include more information about the etiologies of the CP in the children, as previous studies have suggested that there may be differences in cognitive profiles related to subtypes of CP that are not explained by differences in intelligence, but which may be related to the localization of the insult in the brain (Pueyo et al., 2003; van Kampen et al., 2012).

Comparing performance of children with various disabilities and disorders on tasks of the same type as used in this article, including children without motor disabilities and speech impairment, may give information about the role that different forms of language experiences play for planning and other executive functions. It may be useful to substitute some of the easier items within each task with more difficult ones and to compare performance on these items with traditional tests and questionnaires used to measure executive functions.

AUTHOR CONTRIBUTIONS

Kristine Stadskleiv has been involved in collecting, transcribing, coding, and analyzing the data included in the study and has had the main responsibility for preparing the manuscript. Stephen

von Tetzchner is the principal investigator of the study "Becoming an aided communicator" and is responsible for the design of the study, for collecting data from the Norwegian participants, for monitoring the project progress and for preparation of this manuscript. Beata Batorowicz was involved in developing the coding system for the tasks, as well as in analyzing the data and preparing the manuscript. Hans van Balkom, Annika Dahlgren-Sandberg, and Gregor Renner were involved in collecting and analyzing data.

ACKNOWLEDGMENTS

Thanks to all the children, parents, and professionals who participated in the study. The study received funding from the Foundation Sophies Minde, Norway, and from the Department of Psychology, University of Oslo, Norway.

REFERENCES

Alant, E., and Casey, M. (2005). Assessment concessions for learners with impairments. *S. Afr. J. Educ.* 25, 185–189.

Andersen, G. L., Irgens, L. M., Haagaas, I., Skranes, J. S., Meberg, A. E., and Vik, T. (2008). Cerebral palsy in Norway: prevalence, subtypes and severity. *Eur. J. Paediatr. Neurol.* 12, 4–13. doi: 10.1016/j.ejpn.2007.05.001

Andersen, G. L., Mjøen, T. R., and Vik, T. (2010). Prevalence of speech problems and the use of augmentative and alternative communication for children with cerebral palsy: a registry-based study in Norway. *Perspect. Augment. Altern. commun.* 19, 12–20. doi: 10.1044/aac19.1.12

Anderson, P. J. (2008). "Towards a developmental model of executive function," in *Executive Functions and The Frontal Lobes. A Lifespan Perspective*, eds V. Anderson, R. Jacobs, and P. J. Anderson (New York, NY: Taylor and Francis), 3–21.

Astington, J. W., and Hughes, C. (2013). "Theory of mind: self-reflection and social understanding," in *Oxford Handbook of Developmental Psychology*, ed P. D. Zelazo (New York, NY: Oxford University Press), 398–424.

Batorowicz, B., Stadskleiv, K., and von Tetzchner, S. (2013). "Acting on the physical world through aided communication," in *Paper Presented at Regional Conference of the International Society for Early Intervention* (Saint Petersburg).

Batorowicz, B., Stadskleiv, K., and von Tetzchner, S. (2014). "Participation of aided communicators in play with peers: language for action," in *Paper Presented at The 16th Biennial Conference of the International Society for Augmentative and Alternative Communication* (Lisbon).

Beckung, E., White-Koning, M., Marcelli, M., McManus, V., Michelsen, S., Parkes, J., et al. (2008). Health status of children with cerebral palsy living in Europe: a multi-centre study. *Child. Care. Health Dev.* 34, 806–814. doi: 10.1111/j.1365-2214.2008.00877.x

Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50, 7–15. doi: 10.1016/0010-0277(94)90018-3

Benson, J. E., Sabbagh, M. A., Carlsen, S. M., and Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Dev. Psychol.* 49, 1615–1627. doi: 10.1037/a0031056

Bishop, D. V. (2003). *Test for Reception of Grammar – Second Edition (TROG-2)*. Oxford: Pearson Assessment.

Böttcher, L. (2010). Children with spastic cerebral palsy, their cognitive functioning, and social participation. *Child Neuropsychol.* 16, 209–228. doi: 10.1080/09297040903559640

Böttcher, L. (2012). Culture and the learning and cognitive development of children with severe disabilities—continuities and discontinuities with children without disabilities. *Mind Cult. Act.* 19, 89–106. doi: 10.1080/10749039.2011.632050

Böttcher, L., Flachs, E. M., and Uldall, P. (2010). Attentional and executive impairments in children with spastic cerebral palsy. *Dev. Med. Child Neurol.* 52, e42–e47. doi: 10.1111/j.1469-8749.2009.03533.x

Brekke, K. M., and von Tetzchner, S. (2003). "Co-construcion in graphic language development," in *Augmentative and Alternative Communication. Developmental Issues*, eds S. von Tetzchner and N. Grove (London: Whurr/Wiley), 176–210.

Brocki, K. G., and Bohlin, G. (2004). Executive functions in children aged 6 to 13: a dimensional and developmental study. *Dev. Neuropsychol.* 26, 571–593. doi: 10.1037//0012-1649.37.5.715

Caillies, S., Hody, A., and Calmus, A. (2012). Theory of mind and irony comprehension in children with cerebral palsy. *Res. Dev. Disabil.* 33, 1380–1388. doi: 10.1016/j.ridd.2012.03.012

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev. Neuropsychol.* 28, 595–616. doi: 10.1207/s15326942dn2802_3

Clarke, M. T., Loganathan, D., and Swettenham, J. (2012). Assessing true and false belief in young children with cerebral palsy through anticipatory gaze behaviors: a pilot study. *Res. Dev. Disabil.* 33, 2058–2066. doi: 10.1016/j.ridd.2012.05.009

Collins, S. (1996). "Referring expressions in conversations between aided and natural speakers," in *Augmentative and Alternative Communication: European Perspectives,* eds S. von Tetzchner and M. H. Jensen (London: Whurr/Wiley), 89–100.

Collins, S., and Marková, I. (1999). Interactions between impaired and unimpaired speakers: inter-subjectivity and the interplay of culturally shared and situation specific knowledge. *Br. J. Soc. Psychol.* 38, 339–368

Dahlgren Sandberg, A. (2006). Reading and spelling abilities in children with severe speech impairments and cerebral palsy at 6, 9, and 12 years of age in relation to cognitive development: a longitudinal study. *Dev. Med. Child Neurol.* 48, 629–634. doi: 10.1017/S0012162206001344

Dahlgren, S., Dahlgren Sandberg, A. D., and Larsson, M. (2010). Theory of mind in children with severe speech and physical impairments. *Res. Dev. Disabil.* 31, 617–624. doi: 10.1016/j.ridd.2009.12.010

Delis, D. C., Kaplan, E., and Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio, TX: The Psychological Corporation.

Doebel, S., and Zelazo, P. D. (2013). Bottom-up and top-down dynamics in young children's executive function: labels aid 3-year-olds' performance on the Dimensional Change Card Sort. *Cogn. Dev.* 28, 222–232. doi: 10.1016/j.cogdev.2012.12.001

Egeland, J., and Fallmyr, Ø. (2010). Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF): support for a distinction between emotional and behavioral regulation. *Child Neuropsychol.* 16, 326–337. doi: 10.1080/09297041003601462

Eliasson, A. C., Krumlinde Sundholm, L., Rösblad, B., Beckung, E., Arner, M., Öhrvall, et al. (2006). The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev. Med. Child. Neur.* 48, 549–554. doi: 10.1111/j.1469-8749.2006.tb01313.x

Engelhardt, P. E., Nigg, J. T., and Ferreira, F. (2013). Is the fluency of language outputs related to individual differences in intelligence and executive function? *Acta Psychol. (Amst.)*144, 424–432. doi: 10.1016/j.actpsy.2013.08.002

Eysenck, M. W., and Keane, M. T. (1990). *Cognitive Psychology. A Student's Handbook*. Hove: Lawrence Erlbaum Associates.

Falkman, K. W., Dahlgren Sandberg, A., and Hjelmquist, E. (2005). Theory of mind in children with severe speech and physical impairment (SSPI): a longitudinal study. *Int. J. Disabil. Dev. Educ.* 52, 139–157. doi: 10.1080/10349120500086397

Fatzer, S. T., and Roebers, C. M. (2012). Language and executive functions: the effect of articulatory suppression on executive functioning in children. *J. Cogn. Dev.* 13, 454–472. doi: 10.1080/15248372.2011.608322

Ferm, U., Ahlsen, E., and Björck-Åkesson, E. (2005). Conversational topics between a child with complex communication needs and her caregiver at mealtime. *Augment. Altern. Commun.* 21, 19–41. doi: 10.1080/07434610412331270507

Foy, J. G., and Mann, V. A. (2013). Executive function and early reading skills. *Read. Writ.* 26, 453–472. doi: 10.1007/s11145-012-9376-5

Ganesalingam, K., Yeates, K. O., Taylor, H. G., Walz, N. C., Stancin, T., and Wade, S. (2011). Executive functions and social comptence in young children 6 months following traumatic brain injury. *Neuropsychology* 25, 466–476. doi: 10.1037/a0022768

Gerken, L. (2005). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition* 98, B67–B74. doi: 10.1016/j.cognition.2005.03.003

Gioia, G. A., Isquith, P. K., Guy, S. C., and Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function*. Odessa, FL: Psychological Assessment Resources.

Hidecker, M. J. C., Paneth, N., Rosenbaum, P. L., Kent, R. D., Lillie, J., Eulenberg, J. B., et al. (2011). Developing and validating the communication function classification system for individuals with cerebral palsy. *Dev. Med. Child Neurol.* 53, 704–710. doi: 10.1111/j.1469-8749.2011.03996.x

Higginbotham, D. J., Shane, H., Russell, S., and Caves, K. (2007). Access to AAC: present, past, and future. *Augment. Altern. Commun.* 23, 243–257. doi: 10.1080/07434610701571058

Hitch, G. J., and Towse, J. N. (1995). "Working memory: what develops?," in *Memory Performance and Competencies: Issues in Rowth and Development*, eds F. E. Weinert and W. Schneider (Hillsdale: Lawrence Erlbaum Associates), 3–22.

Hudson, K. D., and Farran, E. K. (2011). "Executive function and motor planning," in *Neurodevelopmental Disorders Across the Lifespan: a Neuroconstructivist Approach*, eds E. K. Farran and A. Karmiloff-Smith (Oxford: Oxford University Press), 165–181.

Jenks, K. M., de Moor, J., Lieshout, E. C. D. M., Maathuis, K. G. B., Keus, I., and Gorter, J. W. (2007). The effect of cerebral palsy on arithmetic accuracy is mediated by working memory, intelligence, early numeracy, and instruction time. *Dev. Neuropsychol.* 32, 861–879. doi: 10.1080/87565640701539758

Kinsella, G., Storey, E., and Crawford, J. R. (2007). "Executive function and its assessment," in *Neurology and Clinical Neuroscience,* eds A. H. V Schapira, E. Byrne, S. DiMauro, R. S. J. Frankowiak, R. T. Johnson, Y. Mizuno, et al. (Philadelphia: Mosby Elsevier), 83–95.

Landry, O., Russo, N., Dawkins, T., Zelazo, P. D., and Burack, J. A. (2012). The impact of verbal and nonverbal development on executive function in Down syndrome and Williams syndrome. *J. Dev. Disabl.* 18, 26–35.

Larsson, M., and Dahlgren Sandberg, A. (2008). Memory ability of children with complex communication needs. *Augment. Altern. Commun.* 24, 139–148. doi: 10.1080/07434610801897239

Lezak, M. D. (ed.). (2004). *Neuropsychological Asessment*, 4th Edn. Oxford: Oxford University Press.

Light, J., and Drager, K. (2007). AAC Technologies for young children with complex communication needs: state of the science and future research directions. *Augment. Altern. Commun.* 23, 204–216. doi: 10.1080/07434610701553635

Light, J., and Lindsay, P. (1991). Cognitive science and augmentative and alternative communication. *Augment. Altern. Commun.* 7, 186–203. doi: 10.1080/07434619112331275893

Luria, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behavior*. Oxford: Pergamon Press

Mareschal, D. (2011). From NEOconstructivism to NEUROconstructivism. *Child Dev. Perspect.* 5, 169–170. doi: 10.1111/j.1750-8606.2011.00185.x

McCoy, D. L. C., Raver, C. C., Lowenstein, A. E., and Tirado-Strayer, N. (2011). Assessing self-regulation in the classroom: validation of the BIS-11 and the BRIEF in low-income, ethnic minority school-age children. *Early Educ. Dev.* 22, 883–906, doi: 10.1080/10409289.2010.508371

Miyake, A., Emerson, M. J., Padilla, F., and Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: the effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychol. (Amst.)* 115, 123–142. doi: 10.1016/j.actpsy.2003.12.004

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Murray, J., and Goldbart, J. (2009). Cognitive and language acquisition in typical and aided language learning: a review of recent evidence from an aided communication perspective. *Child Lang. Teach. Ther.* 25, 31–58. doi: 10.1177/0265659008098660

Murray, J., and Goldbart, J. (2011). Emergence of working memory in children using aided communication. *Assist. Technol.* 5, 214–232. doi: 10.1108/17549451111190623

Nicholas, M., Michele, P., Sinotte, M. P., and Helm-Estabrooks, N. (2011). C-Speak Aphasia alternative communication program for people with severe aphasia: importance of executive functioning and semantic knowledge. *Neuropsychol. Rehabil.* 21, 322–366. doi: 10.1080/09602011.2011.559051

Novak, I., McIntyre, S., Morgan, C., Campbell, L., Dark, L., Morton, N., et al. (2013). A systematic review of interventions for children with cerebral palsy: state of the evidence. *Dev. Med. Child Neurol.* 55, 885–910. doi: 10.1111/dmcn.12246

Oxley, J. (2003). "Memory and strategic demands of electronic speech-output communication aids," in *Augmentative and Alternative Communication: Developmental Issues,* eds S. von Tetzchner and N. Grove (London: Whurr/Wiley), 38–66.

Oxley, J., and Norris, J. A. (2000). Children's use of memory strategies: relevance to voice output communication aid use. *Augment. Altern. Commun.* 16, 79–94. doi: 10.1080/07434610012331278924

Palisano, R. J., Rosenbaum, P., Walter, S., Russel, D., Wood, E., and Galuppi, B. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev. Med. Child Neurol.* 39, 214–223. doi: 10.1111/j.1469-8749.1997.tb07414.x

Pennington, L., Mjøen, T., da Graça Andrada, M., and Murray, J. (2010). *Viking Taleskala.* Norway: Sykehuset i Vestfold.

Pirila, S., Meere, J., van der, Rantanen, K., Jokiluoma, M., and Eriksson, K. (2011). Executive function in youth with spastic cerebral palsy. *J. Child Neurol.* 26, 817–821. doi: 10.1177/0883073810392584

Pueyo, R., Junqu,é, C., and Vendrell, P. (2003). Neuropsychological differences between bilateral dyskinetic and spastic cerebral palsy. *J. Child Neurol.* 18, 845–850. doi: 10.1177/088307380301801204

Purdy, M., and Koch, A. (2006). Prediction of strategy usage by adults with aphasia. *Aphasiology* 20, 337–348, doi: 10.1080/02687030500475085

Ratcliff, A. (1994). Comparison of relative demands implicated in direct selection and scanning: considerations from normal children. *Augment. Altern. Commun.* 10, 67–74 doi: 10.1080/07434619412331276770

Sameroff, A. (2010). A unified theory of development: a dialectic integration of nature and nurture. *Child Dev.* 81, 6–22. doi: 10.1111/j.1467-8624.2009.01378.x

Schiørbeck, H., and Stadskleiv, K. (2008). "Utredning og tiltak ved kognitive vansker," in *Habilitering. Tverrfaglig Arbeid for Mennesker Med Utviklingsmessige Funksjonshemninger,* eds S. von Tetzchner, F. Hesselberg, and H. Schiørbeck (Oslo: Gyldendal Akademisk), 239–290.

Sigurdardottir, S., Eiriksdottir, A., Gunnarsdottir, E., Meintema, M., Arnadottir, U., and Vik, T. (2008). Cognitive profile in young Icelandic children with cerebral palsy. *Dev. Med. Child Neurol.* 50, 357–362. doi: 10.1111/j.1469-8749.2008.02046.x

Smith, M. M. (2005). The dual challenges of aided communication and adolescence. *Augment. Altern. Commun.* 21, 67–79. doi: 10.1080/10428190400006625

Strauss, E., Sherman, E. M., and Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. New York, NY: Oxford University Press.

Stuss, D. (1992). Biological and psychological development of executive functions. *Brain Cogn.* 20, 8–23.

Thistle, J. J., and Wilkinson, K. M. (2013). Working memory demands of aided augmentative and alternative communication for individuals with developmental disabilities. *Augment. Altern. Commun.* 29, 235–245. doi: 10.3109/07434618.2013.815800

Usai, M. C., Viterbori, P., Traverson, L., and De Franchis, V. (2013). Latent structure of executive function in five- and six-year-old children: a longitudinal study. *Eur. J. Dev. Psychol.* 11, 447–462. doi: 10.1080/17405629.2013.840578

van Kampen, P. M., Ledebt, A., Smorenburg, A. R. P., Vermeulen, R. J., Kelder, M. E., van der Kamp, J., et al. (2012). Gaze behaviour during interception in children with spastic unilateral cerebral palsy. *Res. Dev. Disabil.* 33, 45–53. doi: 10.1016/j.ridd.2011.08.008

von Tetzchner, S. (1997). The use of graphic language intervention among young children in Norway. *Eur. J. Disord. Commun.* 32, 217–234.

von Tetzchner, S., Almeida, M. A., Balandin, S., Basil, C., Batorowicz, B., Clendon, S., et al. (2012). "The development of aided communicative competence—Communicative problem solving in aided communicators and naturally speaking children aged 5–15 years," in *Paper Presented at the 15th Biennial Conference of the International Society for Augmentative and Alternative Communication*, (Pittsburgh, PA).

von Tetzchner, S., and Basil, C. (2011). Terminology and notation in written representations of conversations with augmentative and alternative communication. *Augment. Altern. Commun.* 27, 141–149. doi: 10.3109/07434618.2011.610356

von Tetzchner, S., and Grove, N. (eds.). (2003). *Augmentative and Alternative Communication. Developmental Issues*. London: Whurr/Wiley

von Tetzchner, S., and Martinsen, H. (1996). "Words and strategies: conversations with young children who use aided language," in *Augmentative and Alternative Communication: European Perspectives* (London: Whurr/Wiley), 65–88.

von Tetzchner, S., and Martinsen, H. (2000). *Introduction to Augmentative and Alternative Communication, 2nd Edn*. London: Whurr/Wiley.

Vygotsky, L. (1986). *Thought and Language*. Cambridge, MA: The MIT Press.

White, D. A., and Christ, S. E. (2005). Executive control of learning and memory in children with bilateral spastic cerebral palsy. *J.Int. Neuropsychol. Soc.* 11, 920–924.

Whittingham, K., Bodimeade, H. L., Lloyd, O., and Boyd, R. N. (2014). Everyday psychological functioning in children with unilateral cerebral palsy: does

executive functioning play a role? *Dev. Med. Child Neurol.* 56, 572–579. doi: 10.1111/dmcn.1237

Wiebe, S. A., Espy, K. A., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575

Willoughby, M., and Blair, C. (2011). Test-retest reliability of a new executive function battery for use in early childhood. *Child Neuropsychol.* 17, 564–579. doi: 10.1080/09297049.2011.554390

Conflict of Interest Statement: The Associate Editors Ayelet Lahat and Oriane Landry declare that, despite being affiliated to the same institution as the author Beata Batorowicz, the review process was handled objectively and no conflict of interest exists. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 January 2014; accepted: 21 August 2014; published online: 08 September 2014.

Citation: Stadskleiv K, von Tetzchner S, Batorowicz B, van Balkom H, Dahlgren-Sandberg A and Renner G (2014) Investigating executive functions in children with severe speech and movement disorders using structured tasks. Front. Psychol. ***5****:992. doi: 10.3389/fpsyg.2014.00992*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Stadskleiv, von Tetzchner, Batorowicz, van Balkom, Dahlgren-Sandberg and Renner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 18 March 2014
doi: 10.3389/fpsyg.2014.00227

Age-related trends of inhibitory control in Stroop-like big–small task in 3 to 12-year-old children and young adults

Yoshifumi Ikeda[1,2] *, Hideyuki Okuzumi[1] and Mitsuru Kokubun[1]

[1]*Department of Special Needs Education, Faculty of Education, Tokyo Gakugei University, Tokyo, Japan*
[2]*Japan Society for the Promotion of Science, Tokyo, Japan*

Edited by:
Yusuke Moriguchi, Joetsu University of Education, Japan

Reviewed by:
Akira Yasumura, National Center of Neurology and Psychiatry, Japan
Ting Xiao, Nanjing Brain Hospital, China
Andrew Simpson, University of Essex, UK

****Correspondence:***
Yoshifumi Ikeda, Department of Special Needs Education, Faculty of Education, Tokyo Gakugei University, 4-1-1 Nukuikita, Koganei, Tokyo 184-8501, Japan
e-mail: r113001n@st.u-gakugei.ac.jp

Inhibitory control is the ability to suppress competing, dominant, automatic, or prepotent cognitive processing at perceptual, intermediate, and output stages. Inhibitory control is a key cognitive function of typical and atypical child development. This study examined age-related trends of Stroop-like interference in 3 to 12-year-old children and young adults by administration of a computerized Stroop-like big–small task with reduced working memory demand. This task used a set of pictures displaying a big and small circle in black and included the same condition and the opposite condition. In the same condition, each participant was instructed to say "big" when viewing the big circle and to say "small" when viewing the small circle. In the opposite condition, each participant was instructed to say "small" when viewing the big circle and to say "big" when viewing the small circle. The opposite condition required participants to inhibit the prepotent response of saying the same, a familiar response to a perceptual stimulus. The results of this study showed that Stroop-like interference decreased markedly in children in terms of error rates and correct response time. There was no deterioration of performance occurring between the early trials and the late trials in the sessions of the day–night task. Moreover, pretest failure rate was relatively low in this study. The Stroop-like big–small task is a useful tool to assess the development of inhibitory control in young children in that the task is easy to understand and has small working memory demand.

Keywords: inhibition, executive function, cognitive control, day–night task, Stroop

INTRODUCTION

Inhibitory control is the ability to suppress competing, dominant, automatic, or prepotent cognitive processing at perceptual, intermediate, and output stages (Nigg, 2000; Friedman and Miyake, 2004; see Ikeda et al., 2013a, for a discussion of classification in inhibitory control). That ability is used when the cognitive processing must be suppressed merely because it is inappropriate and when the cognitive processing must be suppressed in favor of a subdominant but appropriate one. Inhibitory control has been suggested as playing a critical role in executive function: higher order cognitive function that coordinates a goal-directed behavior (Harnishfeger and Bjorklund, 1994; Miyake et al., 2000; Anderson, 2001, 2002; Miyake and Friedman, 2012). Deficits in inhibitory control have also been implicated in behavioral problems associated with developmental disorders such as attention deficit hyperactivity disorder (ADHD) (Barkley, 1997; Ozonoff and Jensen, 1999; Friedman and Miyake, 2004; Spronk et al., 2008; Song and Hakoda, 2011; Yasumura et al., 2014). Inhibitory control is a key cognitive function of typical and atypical child development.

Inhibitory control is often measured in young children by administering the Stroop-like day–night task (Gerstadt et al., 1994; for a review, see Montgomery and Koeltzow, 2010). In the day–night task, children are presented with either a day picture of the sun or a night picture of the moon and stars, and they are instructed to say "day" to the night picture and "night" to the day picture. During the task, children must (a) suppress a dominant response of naming what a picture represents and (b) execute a competing subdominant response based on the instructions. Previous reports have described that performance of the day–night task improves significantly in young children (Montgomery and Koeltzow, 2010). Gerstadt et al. (1994) reported that the Stroop-like interference, measured as the difference of response time (RT) between experimental (saying the opposite of what is shown for day/night cards) and control conditions (saying "day" or "night" to abstract shapes), decreases in children between ages 3.5 and 5. Accuracy also improves concomitantly with age in children aged 3–7. Recent studies confirmed these findings, using variants of the day–night task with a range of stimuli and responses, such as color labels and basic-level object names (e.g., Simpson and Riggs, 2005a,b).

Difficulty in the day–night paradigm is believed to arise because of response competition occurring within the response set during testing. Although it was expected that the stronger association between word pairs makes the task more inhibitory demanding, recent research has demonstrated that what causes prepotency of response is not the relation between the response-to-be activated and the response-to-be suppressed (Diamond et al., 2002) but membership in the response set (Simpson and Riggs, 2005a). The problem is that the correct response on one trial is also the incorrect but prepotent response on subsequent trials. The

potency of the incorrect response is magnified on each trial because of its activation during testing (by virtue of its inclusion in the response set) coupled with its depiction throughout the testing. In other words, the structure of the day–night task elevates response competition because (a) the response alternative that must be suppressed is depicted on the test card (e.g., "night" for the night card) and (b) the incorrect response alternative was previously activated on previous trials (e.g., "night" for day cards) in the case of a correct response, i.e., a response set effect (Simpson and Riggs, 2005a; Montgomery et al., 2008; Simpson et al., 2012).

Working memory is also presumed to be an important factor related to the performance in the day–night paradigm. Working memory may be involved because resolving which of the conflicting responses must be suppressed entails holding the task rules in mind ("say 'day' for night card" and "say 'night' for day card"). In fact, some studies report deterioration in performance occurring between the first four trials and the last four trials in the sessions of the day–night task in young children (e.g., Gerstadt et al., 1994). These reports suggest that young children may have forgotten rules or that working memory demands add to the processing requirements of the task and consequently compromise inhibitory mechanisms (Montgomery and Koeltzow, 2010). Then, it is expected to reduce the working memory demands in the day–night paradigm so that the task primarily measures inhibitory control.

Working memory demands may also be related to learning the task rules. It might be true that the day–night task recruits working memory to a certain extent because learning the combination between words and pictures ("day" for a night picture of the moon and stars and "night" for a day picture of the sun) is not easy to understand for young children. Actually, previous studies with the day–night task had a great pretest failure rate, especially in children aged between 3 and 4 (e.g., Gerstadt et al., 1994). A problem is that with more children failing the pretest, the sample of children whose RT and accuracy data are analyzed is not representative of the population. Probably, the children with the poorest inhibitory control get excluded.

This study was conducted to examine age-related trends of inhibitory control in various age groups by administration of a Stroop-like day–night variant with reduced working memory demand. For this study, 3 to 12-year-old children and young adults were administered a Stroop-like big–small task. The task used a set of pictures displaying a big or small circle in black and required participants to produce sized-based responses following instructions given by the experimenter. The size labels "big and small" were used because they are well understood even by very young children and because they are distinctive and opposite, both of which may facilitate learning and holding the task rules. The task has two conditions: the opposite condition, in which a participant says the opposite of what is shown with card pairs, and the same task condition, in which a participant simply names what the stimulus represents. Because the original study (Gerstadt et al., 1994) used a different combination of words and pictures in the opposite condition (saying the opposite of what is shown for day–night cards) to that used in the control condition (saying "day" or "night" to abstract shapes), the degree to which a picture evokes a particular response was not controlled. Comparison between opposite and same conditions, as in this study, indicates the inhibitory processes, controlling for a difference in the degree to which a picture evokes a particular response. In this study, unlike the standard "card" version of the day–night task, the task was computerized to evaluate the correct RT more precisely.

MATERIALS AND METHODS

PARTICIPANTS

Participants were 113 typically developing people who were divided into six age groups: (a) 3–4 year, 20 children (10 boys, 10 girls; M age = 51.5 months, age range = 43–59); (b) 5–6 year, 14 children (7 boys, 7 girls; M age = 68.6 months, age range = 60–83); (c) 7–8 year, 20 children (11 boys, 9 girls; M age = 95.1 months, age range = 87–107); (d) 9–10 year, 19 children (9 boys, 10 girls; M age = 119.6 months, age range = 108–131); (e) 11–12 year, 17 children (9 boys, 8 girls; M age = 144.4 months, age range = 133–153); and (f) 23 young adults (9 men, 14 women; M age = 21.1 year, age range = 18–24). Children were recruited through local mainstream preschool and elementary school programs. Young adults were recruited from a university. All participants speak Japanese as a first language. Criteria for inclusion were the absence of bilingualism and absence of behavioral or educational problems, which would affect the study of inhibitory control.

MATERIALS AND METHODS

The Stroop-like big–small task used a set of pictures displaying either a big (12 cm diameter) or a small (1 cm diameter) circle in black. The same set of pictures was used in the same and opposite conditions. SuperLab (Cedrus Corp., San Pedro, CA, USA) controlled the task, presenting stimuli and recording participants' vocal responses (error and time).

PROCEDURES

Participants were tested individually in quiet rooms at their respective schools. At arrival, a participant was asked to be seated next to the experimenter at the table and approximately 50 cm in front of a monitor with a headset microphone. Subsequently, the experimenter explained that they were going to play a "game" in which they would see two pictures. The experimenter showed the participant the big and small circles at the same time on the screen and asked him or her to point to the big circle and the small circle in turn. All participants were able to do this. Then, each was administered the Stroop-like big–small task. In this task, the same condition was arranged to precede the opposite condition in an attempt to elicit robust interference.

Prior to the test phase, participants were trained on how to play each "game." For the same condition, the experimenter said, "Here is a picture of big circle (show a big circle on the monitor). When this picture is shown, I want you to say 'big'. And, here is a picture of small circle (show a small circle on the monitor). When this picture is shown, I want you to say 'small'." The participant did four practice trials (big–small–small–big). If the participant made any error, then the participant was corrected, reminded of the rules, and administered another four practice trials. The task did not commence until the participant was 100% correct for a set. For the opposite condition,

the experimenter said, "We are going to play an opposite game. Here is a picture of big circle (show a big circle on the monitor). When this picture is shown, I want you to say 'small'. And, here is a picture of small circle (show a small circle on the monitor). When this picture is shown, I want you to say 'big'." The opposite condition required participants to inhibit the prepotent response of saying the same, a familiar response to a perceptual stimulus. The practice trials were identical to the same condition.

During the test phase, the participant was asked to respond as quickly and accurately as possible to a series of 20 stimuli (10 big circles and 10 small circles) for each task condition. All stimuli were presented one at a time and randomly at the center of the white screen on the monitor. At the instant a participant's voice key was input, each stimulus was replaced by a fixation cross until the participant was judged by the experimenter to be ready to proceed to the next trial, looking at the fixation cross. The interstimulus interval was not controlled by SuperLab, as it was in a card version of the task, because some younger children have difficulty engaging in the task continuously. No feedback reminding participants of the task rules was given during testing.

ANALYSIS

Numbers of errors and RT for correct responses were recorded. Trials were counted as incorrect when participants' initial responses were errors, even if they self-corrected. The RT was measured as the interval in milliseconds between the presentation of a stimulus and the onset of the participant's vocal response by the microphone. Analysis of RT was conducted only for the correct response. Mean and standard deviations of error rates and correct RT on the whole trials were calculated for each task condition. To examine changes of performance over the course of a session for each task condition, mean and standard deviations of error rates and correct RT were calculated for the first five trials and the last five trials, respectively.

The data were analyzed using analysis of variance (ANOVA). Specifically, two-way ANOVAs with the within-participant factor of condition (same and opposite) and between-participant factor of age group (3 to 4-year olds, 5 to 6-year olds, 7 to 8-year olds, 9 to 10-year olds, 11 to 12-year olds, and young adults) were conducted for error rates and correct RT. Also, three-way ANOVAs with the within-participant factors of condition (same and opposite) and serial position (first five trials and last five trials) and between-participant factor of age group (3 to 4-year olds, 5 to 6-year olds, 7 to 8-year olds, 9 to 10-year olds, 11 to 12-year olds, and young adults) were conducted for error rates and correct RT. Software was used for statistical analyses (SPSS 19.0 for Windows; SPSS Japan Inc., Tokyo, Japan). Unless otherwise noted, a 0.05 level of significance was adopted for all statistical analyses.

ETHICAL APPROVAL

Informed consent was obtained from all adult participants and from a parent of each child participant before the assessment session. Our experimental protocol was administered in accordance with the guidelines of the Declaration of Helsinki and was approved by the institutional review board.

RESULTS

TREATMENT OF UNUSED DATA

An additional 14 participants were tested. Data from 9 participants were not included in this study because they showed results more than 3 SD from the mean of each age group (i.e., outliers). One 3-year-old child was not able to pass a pretest for the saying-opposite condition. Another 3-year-old child was not able to complete the task because his voice was too small to record. Three school-age children were excluded because of experimental error in recording the data.

ACCURACY OF RESPONSE

Figure 1 depicts the mean and standard deviations for the error rates in the Stroop-like big–small task by age group and condition. A 6 (age group) × 2 (condition) mixed-model ANOVA was conducted of the error rates. The analysis showed significant main effects for the age group ($F_{5,107} = 18.40$, $p < 0.001$, partial $\eta^2 = 0.46$), for condition ($F_{1,107} = 26.73$, $p < 0.001$, partial $\eta^2 = 0.20$), and for interaction of age group and condition ($F_{5,107} = 4.02$, $p < 0.01$, partial $\eta^2 = 0.16$). *Post hoc* Bonferroni tests yielded significant differences between-age group comparisons between 3 to 4-year olds and other age groups and between 5 to 6-year olds and other age groups for each condition ($p < 0.05$). For each condition, between-age group comparisons among 7 to 8-year olds, 9 to 10-year olds, and 11 to 12-year olds and young adults were not significant. *Post hoc* Bonferroni tests also yielded significant differences between conditions for 3 to 4-year olds and 5 to 6-year olds ($p < 0.05$), but not for other age groups (7 to 8-year olds, $p = 0.084$; 9 to 10-year olds, $p = 0.107$; 11 to 12-year olds, $p = 0.588$; young adults, $p = 1.00$). These results clarified that the interaction between age group and condition reflected age-related convergence of error rates in the same condition and in the opposite condition.

Figure 2 depicts the mean and standard deviations for the error rates early (the first five trials) and late (the last five trials) in the session in the Stroop-like big–small task by age group and condition. A 6 (age group) × 2 (condition) × 2 (serial position) mixed-model ANOVA was conducted of the error rates. The analysis showed significant main effects for the age group ($F_{5,107} = 18.33$, $p < 0.001$,

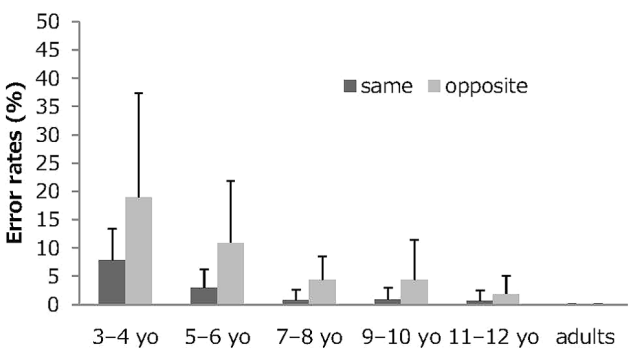


FIGURE 1 | Mean error rates in the Stroop-like big–small. Error bars represent standard deviations. No adults showed an error.

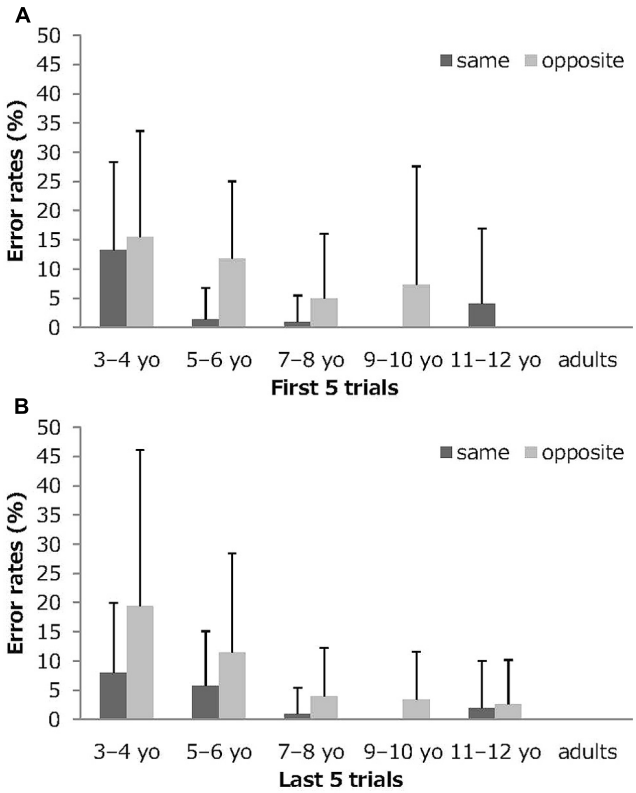


FIGURE 2 | Mean error rates early and late in the session in the Stroop-like big–small. (A) mean error rates on the first five trials, **(B)** mean error rates on the last five trials. Error bars represent standard deviations. On the first five trials, no error was shown for 9 to 10-year olds in the same condition, for 11 to 12-year olds in the opposite condition, and for adults in the same and opposite conditions. On the last five trials, no error was shown for 9 to 10-year olds in the same condition and for adults in the same and opposite conditions.

partial $\eta^2 = 0.46$), for condition ($F_{1,107} = 15.97$, $p < 0.001$, partial $\eta^2 = 0.13$), and for interaction of age group and condition ($F_{5,107} = 2.84$, $p < 0.05$, partial $\eta^2 = 0.12$). Main effect for the serial position, the other two-way interactions, and the three-way interaction were not significant. These results clarified that there were no deterioration of performance over the course of a session in terms of the error rates.

CORRECT RESPONSE TIME

Figure 3 depicts the mean and standard deviations for the correct RTs in the Stroop-like big–small task by age group and condition. A 6 (age group) × 2 (condition) mixed-model ANOVA was conducted of the correct RT. The analysis showed significant main effects for age group ($F_{5,107} = 53.77$, $p < 0.001$, partial $\eta^2 = 0.72$), for condition ($F_{1,107} = 203.02$, $p < 0.001$, partial $\eta^2 = 0.66$), and for interaction of the age group and condition ($F_{5,107} = 34.37$, $p < 0.001$; partial $\eta^2 = 0.62$). *Post hoc* Bonferroni tests yielded all significant between-age group comparisons ($p < 0.05$), except for between 7 to 8-year olds and 9 to 10-year olds and between 9 to 10-year olds and 11 to 12-year olds, for each condition. *Post hoc* Bonferroni tests also yielded significant differences between conditions for 3 to 4-year olds, 5 to 6-year olds, 7 to 8-year olds, and 9 to 10-year olds ($p < 0.05$), but not for other age groups. Results clarified that the interaction between age group and condition reflected age-related convergence of error rates in the same condition and in the opposite condition.

Figure 4 depicts the mean and standard deviations for the correct RTs early (the first five trials) and late (the last five trials) in the session in the Stroop-like big–small task by age group and condition. A 6 (age group) × 2 (condition) × 2 (serial position) mixed-model ANOVA was conducted of the correct RT. The analysis showed significant main effects for the age group ($F_{5,107} = 46.25$, $p < 0.001$, partial $\eta^2 = 0.68$), for condition ($F_{1,107} = 153.51$, $p < 0.001$, partial $\eta^2 = 0.59$), and for interaction of age group and condition ($F_{5,107} = 25.24$, $p < 0.001$, partial $\eta^2 = 0.54$). Main effect for the serial position, the other two-way interactions, and the three-way interaction were not significant. These results clarified that there were no deterioration of performance over the course of a session in terms of the correct RTs.

DISCUSSION

This study examined age-related trends of Stroop-like interference in 3 to 12-year-old children and young adults by administration of a computerized Stroop-like big–small task. In this study, the differences between the opposite and same conditions were compared among age groups for error rates and correct RT. It was hypothesized that working memory demand is reduced in the Stroop-like big–small task.

Results show that Stroop-like interference decreased markedly in children. The difference between conditions in error rates was significant for 3 to 4-year olds and 5 to 6-year olds but not for the older age groups although there were trends toward significance for some older age groups, which may be due to relatively small sample size. These results are consistent with the results obtained from previous studies using the day–night task and other variants of this task (Gerstadt et al., 1994; Simpson and Riggs, 2005b). However, this difference in correct RT was significant for 3 to 4-year olds, 5 to 6-year olds, 7 to 8-year olds, and 9 to 10-year olds. This result is consistent with the results reported by Simpson and

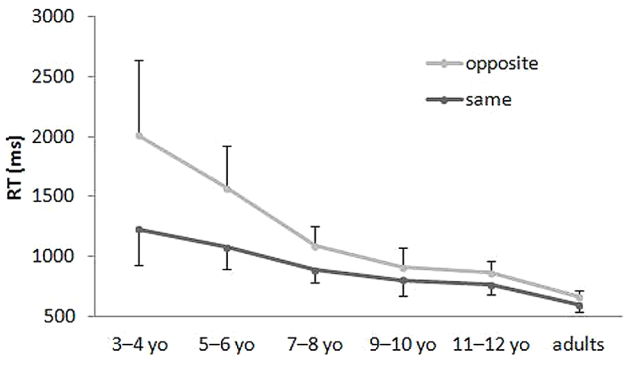


FIGURE 3 | Mean RT in the Stroop-like big–small. Error bars represent standard deviations.

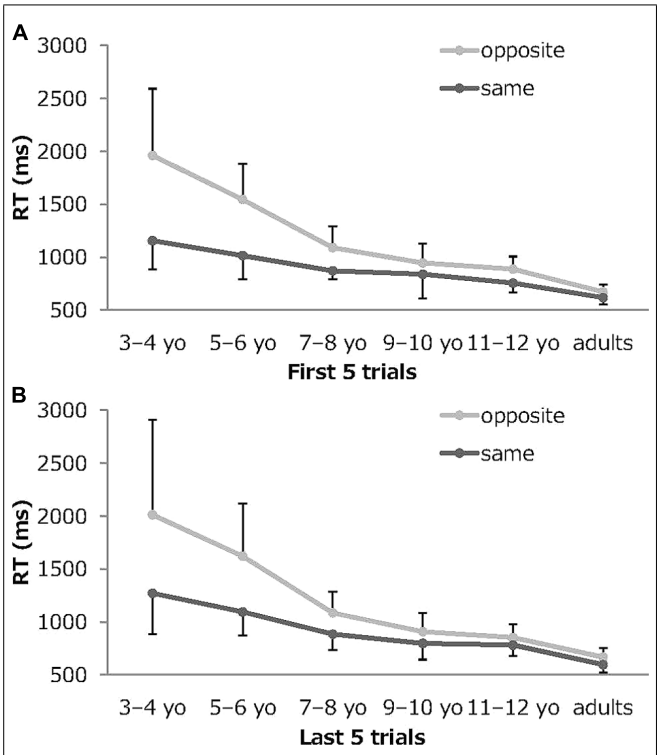


FIGURE 4 | Mean RT early and late in the session in the Stroop-like big–small. (A) mean RT on the first five trials, **(B)** mean RT on the last five trials. Error bars represent standard deviations.

Riggs (2005b), which used five age groups (3.5-, 5-, 7-, 9-, and 11-year olds) and reported that Stroop-like interference was greatest in 3.5-year olds, greatly reduced in 5-year olds, and thereafter declined more moderately up to the age of 11. In this study, the difference between 9 to 10-year olds and young adults were not significant, although a decrement of interference between older children and young adulthood was often observed in the Stroop color-word task (Ikeda et al., 2011, 2013b) and the Stroop-like task (Ikeda et al., 2013a). This decrement can be interpreted as reflecting reduced inhibitory demand in the Stroop-like big–small task compared to other inhibitory tasks.

This study used a variant of the day–night task particularly addressing the concept of size, "big" and "small." These sizes were concrete for participants in this study because they were perceptual features of the stimuli that were used, which seemed to facilitate sampling of young children, having them feel more comfortable by learning and holding the rules in mind. Actually, fewer children refused participation or were unable to pass the pretest, compared to those of the original study using the day–night task (Gerstadt et al., 1994). A problem for previous research is that with more children failing the pretest, the sample of children whose RT and accuracy data are analyzed is not representative of the population. Moreover, the results showed no difference in error rates and correct RTs between the first five trials and the last five trials in the session, suggesting that participants did not forget the rules or that working memory recruited in the task did not compromise inhibitory mechanisms.

In conclusion, this study demonstrated that the difference between naming what stimuli represent and naming of the "opposite" of the stimuli was decreased significantly during young childhood in the Stroop-like big–small task that has smaller working memory demands than the original version of the day–night task. In other words, this study showed that inhibitory control develops rapidly in young children. The Stroop-like big–small task is a useful tool to investigate the development of inhibitory control in young children in that the task is easy to understand and has small working memory demand.

ACKNOWLEDGMENTS

The authors thank all who participated in the study. This research was supported by the Japan Society for the Promotion of Science Research Fellowship for Young Scientists (to Yoshifumi Ikeda).

REFERENCES

Anderson, M. (2001). Annotation: conceptions of intelligence. *J. Child Psychol. Psychiatry* 42, 287–298. doi: 10.111/1469-7610.00722

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* 121, 65–94. doi: 10.1037/0033-2909.121.1.65

Diamond, A., Kirkham, N., and Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Dev. Psychol.* 38, 352–362. doi: 10.1037/0012-1649.38.3.352

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135. doi: 10.1037/0096-3445.133.1.101

Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action: performance of children 3 1/2–7 years old on a Stroop-like day–night test. *Cognition* 53, 129–153. doi: 10.1016/0010-0277(94)90068-X

Harnishfeger, K. K., and Bjorklund, D. F. (1994). A developmental perspective on individual differences in inhibition. *Learn. Individ. Differ.* 6, 331–355. doi: 10.1016/1041-6080(94)90021-3

Ikeda, Y., Okuzumi, H., and Kokubun, M. (2013a). Age-related trends of Stroop-like interference in animal size tests in 5–12-year-old children and young adults. *Child Neuropsychol.* 19, 276–291. doi: 10.1080/09297049.2012.658364

Ikeda, Y., Okuzumi, H., and Kokubun, M. (2013b). Stroop/reverse-Stroop interference in typical development and its relation to symptoms of ADHD. *Res. Dev. Disabil.* 34, 2391–2398. doi: 10.1016/j.ridd.2013.04.019

Ikeda, Y., Okuzumi, H., Kokubun, M., and Haishi, K. (2011). Age-related trends of interference control in school-age children and young adults in the Stroop color-word test. *Psychol. Rep.* 108, 577–584. doi: 10.2466/04.10.22.PR0.108.2.577-584

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Montgomery, D. E., Anderson, M., and Uhl, E. (2008). Interference control in preschoolers: factors influencing performance on the day–night task. *Infant Child Dev.* 17, 457–470. doi: 10.1002/icd.559

Montgomery, D. E., and Koeltzow, T. E. (2010). A review of the day–night task: the Stroop paradigm and interference control in young children. *Dev. Rev.* 30, 308–330. doi: 10.1016/j.dr.2010.07.001

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychol. Bull.* 126, 220–246. doi: 10.1037/0033-2909.126.2.220

Ozonoff, S., and Jensen, J. (1999). Brief report: specific executive function profiles in three neurodevelopmental disorders. *J. Autism Dev.* 29, 171–177. doi: 10.1023/A:1023052913110

Simpson, A., and Riggs, K. J. (2005a). Factors responsible for performance on the day–night task: response set or semantics? *Dev. Sci.* 8, 360–371. doi: 10.1111/j.1467-7687.2005.00424.x

Simpson, A., and Riggs, K. J. (2005b). Inhibitory and working memory demands of the day–night task in children. *Br. J. Dev. Psychol.* 23, 471–486. doi: 10.1348/026151005X28712

Simpson, A., Riggs, K. J., Beck, S. R., Gorniak, S. L., Wu, Y., Abbott, D. et al. (2012). Refining the understanding of inhibitory processes: how response prepotency is created and overcome. *Dev. Sci.* 15, 62–73. doi: 10.1111/j.1467-7687.2011.01105.x

Song, Y., and Hakoda, Y. (2011). An asymmetric Stroop/reverse-Stroop interference phenomenon in ADHD. *J. Atten. Disord.* 15, 499–505. doi: 10.1177/1087054710367607

Spronk, M., Jonkman, L. M., and Kemner, C. (2008). Response inhibition and attention processing in 5- to 7-year-old children with and without symptoms of ADHD: an ERP study. *Clin. Neurophysiol.* 119, 2738–2752. doi: 10.1016/j.clinph.2008.09.010

Yasumura, A., Kokubo, N., Yamamoto, H., Yasumura, Y., Nakagawa, E., Kaga, M., et al. (2014). Neurobehavioral and hemodynamic evaluation of Stroop and reverse Stroop interference in children with attention-deficit/hyperactivity disorder. *Brain* Dev. 36, 97–106. doi: 10.1016/j.braindev.2013.01.005

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 November 2013; accepted: 28 February 2014; published online: 18 March 2014.

Citation: Ikeda Y, Okuzumi H and Kokubun M (2014) Age-related trends of inhibitory control in Stroop-like big–small task in 3 to 12-year-old children and young adults. Front. Psychol. ***5****:227. doi: 10.3389/fpsyg.2014.00227*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Ikeda, Okuzumi and Kokubun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 05 March 2015
doi: 10.3389/fpsyg.2014.01470

Age differences in high frequency phasic heart rate variability and performance response to increased executive function load in three executive function tasks

Dana L. Byrd[1]*, Erin T. Reuther[2], Joseph P. H. McNamara[3], Teri L. DeLucca[4] and William K. Berg[3]

[1] *Psychology and Sociology, Texas A&M University-Kingsville, Kingsville, TX, USA*
[2] *Department of Psychiatry, LSU Health Sciences Center-New Orleans, New Orleans, LA, USA*
[3] *University of Florida, Gainesville, FL, USA*
[4] *Nemours BrightStart, Jacksonville, FL, USA*

Edited by:
Philip D. Zelazo, University of Minnesota, USA

Reviewed by:
Peter J. Anderson, Murdoch Childrens Research Institute, Australia
Josef Martin Unterrainer, Universitätsmedizin der Johannes Gutenberg – Universität Mainz/Klinik und Poliklinik für Psychosomatische Medizin und Psychotherapie/ Medizinische Psychologie und Medizinische Soziologie, Germany

****Correspondence:***
Dana L. Byrd, Psychology and Sociology, Texas A&M University-Kingsville, MSC177 Station 1, Kingsville, TX, USA
e-mail: dana.byrd.phd@gmail.com

The current study examines similarity or disparity of a frontally mediated physiological response of mental effort among multiple executive functioning tasks between children and adults. Task performance and phasic heart rate variability (HRV) were recorded in children (6 to 10 years old) and adults in an examination of age differences in executive functioning skills during periods of increased demand. Executive load levels were varied by increasing the difficulty levels of three executive functioning tasks: inhibition (IN), working memory (WM), and planning/problem solving (PL). Behavioral performance decreased in all tasks with increased executive demand in both children and adults. Adults' phasic high frequency HRV was suppressed during the management of increased IN and WM load. Children's phasic HRV was suppressed during the management of moderate WM load. HRV was not suppressed during either children's or adults' increasing load during the PL task. High frequency phasic HRV may be most sensitive to executive function tasks that have a time-response pressure, and simply requiring performance on a self-paced task requiring frontal lobe activation may not be enough to generate HRV responsitivity to increasing demand.

Keywords: planning, inhibition (psychology), working memory, heart rate variability, respiratory sinus arrhythmia, child, adult

INTRODUCTION

Executive function is an umbrella term used to group a variety of complex cognitive functions that utilize the attentional control unit of Baddeley's working memory (WM) model which governs allocation of attention and inhibition of automatic or incorrect action. This Central Executive of Baddeley's model utilizes neural connections within the frontal lobes as part of their neural circuitry (Baddeley, 1996; Banich et al., 2000; Jansma et al., 2000; Newman et al., 2003; Owen et al., 2005). This category of executive functions includes a number of abilities and their related tasks. A latent factor analysis of performance on a large number of executive tasks found both a unity to executive functions, as well as separate categories of executive functions (Miyake et al., 2000). For both adults and children, the separate categories included updating of WM and inhibition of automatic/over-learned responses, as well as shifting of attention and action (Miyake et al., 2000; Huizinga et al., 2006). Another executive function, multistep planning toward a goal, has been found to rely on attentional control (Baddeley, 1996) and frontal lobe functioning (Luria, 1966; Shallice, 1982; Unterrainer et al., 2004a; Kaller et al., 2011).

There is a prolonged child development of neural circuitry that differs for various executive functions shows increases in area growth, efficiency of activity, and myelination including in frontal areas from preschool to late adolescence, as well as increased coordination with age of frontal connections' coordinated neural functioning as measured by electroencephalographic coherence (e.g., Casey, 1992; Casey et al., 1997; Hanlon et al., 1999; Thomas et al., 1999; Nelson et al., 2000; Durston et al., 2001; Fuster, 2002; Thatcher et al., 2008). As might be expected, this is accompanied by a prolonged development of executive function task performance, with particularly large improvements during preschool/kindergarten and adolescent years (Klahr and Robinson, 1981; Welsh et al., 1991; Diamond and Taylor, 1996; Denot-Ledunois et al., 1998; Zelazo, 2000; Davidson et al., 2006; Lamm et al., 2006; Zelazo and Müller, 2007; Kaller et al., 2008; Unterrainer et al., 2014).

PHASIC HIGH FREQUENCY HEART RATE VARIABILITY

A psychophysiological measure, phasic high frequency heart rate variability (HRV), may provide valuable information about the modulation of executive control in children and adults. According to Thayer et al. (2009), prior to 1867 Claude Bernard was the first to suggest that cortical activity has a reactive response on heart rate. Since then, it has been found that the heart rate can fluctuate at a wide range of frequencies slow, medium, and fast (Jennings and Yovetich, 1991), with the faster frequency associated with typical inhalation and exhalation rates. Thus, respiratory related HRV has been measured as the spectral power of the heart rate changes within the frequency range of respiration. This measure is somewhat similar to another psychophysiological measure, respiratory sinus arrhythmia (RSA), also measures the synchrony of respiration and heart rate.

Although different terms have been used to define the cognitive process indexed by high frequency HRV, referred to from this point on as HRV, or RSA, there is general agreement that reduced/suppressed RSA or HRV power (less power in the frequency band of respiration) is associated with increased effortful mental processing or effortful attentional control in adults (Porges and Byrne, 1992; Suess et al., 1994; Berntson et al., 1997; Beauchaine, 2001; Hansen et al., 2003; Mulder et al., 2004; Porges, 2007). While it has been concluded that HRV is related to executive task performance and reflects the prefrontal utilization required by active control of attention, there is also a call for further research into which executive functions change HRV (Thayer et al., 2009). This report was limited to examining primarily inhibitory processes.

Empirical evidence supports this link between frontal lobe activation and mental effort during executive function. The role of the frontal cortex in the regulation of HRV has been demonstrated with clinical populations (Althaus et al., 1999, 2004; Lane et al., 2001) as well as functional imaging studies with normative populations (Gianaros et al., 2004; Matthews et al., 2004).

A number of functional imaging – HRV studies have found a relationship between increased activation of the anterior cingulate cortex (ACC) and decreased RSA in frequencies similar to that of respiration (Critchley et al., 2003; Matthews et al., 2004). It has been theorized that the ACC serves to detect instances where it is necessary to recruit frontal areas, including the dorsolateral prefrontal cortex, to manage increasing executive demands (Botvinick et al., 2001; Hajcak et al., 2003).

In a model of the heart–brain connection by Thayer et al. (2009), sympathetic and parasympathetic regulation of HRV is modeled as modulating with increased dorsolateral prefrontal and ACC activation such that increased activation results in decreased HRV. Additionally, part of a model by Thayer et al. (2009) suggest that activation of the prefrontal cortex can result in discontrol of the heart rate response through both a tonic acceleratory drive and a tonic deceleratory drive from both the sympathetic and parasympathetic branches of the autonomic nervous system. We suggest that this results in disregulation of the heart rate response, which we propose would result in decreased phasic high frequency HRV.

Due to their undeveloped frontal neural circuitry, children may be less able or less consistent in their ability to activate the ACC and recruit their underdeveloped frontal areas to manage the difficult executive task conditions, thus deregulating their HRV. That is, children may have less ability to fully recruit the attentional/behavioral control system, including the dorsolateral prefrontal cortex, in order to manage the task conditions. However, this stands in opposition to a model by Thayer et al. (2009) which suggests less activation of the prefrontal cortex would lead to activation of the central nucleus of the amygdala, which would lead to an increase in sympathetic activity and inhibition of the parasympathethoexcitatory neurons, which in turn would lead to a decrease in vagal tone and HRV. It is worth note that this model is based on animal models and adult neurology, and may not apply to the hypofrontality due to a lack of development.

THE CURRENT STUDY

As compared to HRV during a rest period, decreases in HRV have been found during executive function tasks with both adults (Hansen et al., 2003; Johnsen et al., 2003) and children (Hickey et al., 1995, Mezzacappa et al., 1998). We expect from prior research that children will show less HRV responsitivity during the Stroop task (as seen in a younger and older adult developmental study of a variant of the Stroop task, Mathewson et al., 2010), and perhaps also the Tower of London task that requires inhibition of inefficient moves to make the correct counterintuitive correct moves. Studies of HRV during executive function tasks have not, however, been simultaneously assessed with multiple subtypes of executive functions, especially in children. One study compared HRV during Stroop task and mental arithmetic in older adults. This study found that mentally stimulating activities predicted HF-HRV (Lin et al., 2013). However, this was one formal executive function task, the Stroop, and another cognitively challenging task, which likely requires executive functions such as WM, mental arithmetic. The current study, a study including tasks tapping into the subtypes of executive functions would allow for the comparison of developmental differences, which developmental studies show less HRV response in children and animal models and adult neurology suggest increased HRV. We expect to see both performance on each dimension of executive function and in the HRV changes that are associated with increased executive functioning load in both age groups, but our hypotheses about developmental HRV are exploratory.

In the current study we have the goal to examine our child group for developmental differences compared to adults. We approach these goals using three executive functioning tasks which typify subtypes of executive function (Baddeley, 1996; Miyake et al., 2000): inhibition of an automatic/over-learned response, goal-focused multi-step planning, and WM updating. With these tasks we utilized a parametric study design rather than a baseline rest design. A parametric design allows for the calculation of difference scores to a low level of the task to assess increases or decreases in adults' and children's physiological and behavioral responsitivity to increased executive functioning load without confounds possible due to individual or developmental differences in interpretation/processing of a rest baseline. In fact we suggest that a rest baseline may be inappropriate for HRV as it is for other psychophysiological measures such as electroencephalogram and functional magnetic resonance imaging as the baseline of rest requires a form of mental effort, especially in children, as they exhibit attentional and motor control and "tune out" all modalities and inhibit all behavioral responses, which may be a challenge when they are in the novel laboratory environment with electrodes and a respiration band on their bodies.

Functional magnetic resonance imaging suggests that a rest baseline is not a "zero" (Stark and Squire, 2001). A parametric design, rather than a rest baseline, is now becoming standard in functional magnetic resonance imaging, especially in developmental studies (Katsoni et al., 2006). Scores, both correctness and speed, were calculated as difference scores relative to the easiest condition. These difference scores allow for the assessment of the participants' behavioral and physiological reaction to increased cognitive load while controlling for stimulus and developmental

motor events. Though the easiest condition of each task may have some executive load, our parametric design still examines changes in performance and physiological response from a lower level of executive load to higher levels of executive load. HRV responsitivity scores and behavioral responsitivity scores are both calculated similarly to task-rest baseline difference scores but are instead calculated relative to the within-task lowest cognitive demand condition. Our primary hypotheses concerned changes due to task difficulty, and thus these scores reflect the response to the increased task demand.

We examined the executive functions of WM and inhibition as they have been found to be a separate factor (Huizinga et al., 2006) and also planning as it requires the combination of WM and inhibition, as well as longer term goal tactic, and is also crucial task for daily functioning (Luria, 1966). We contrasted adult responses with early elementary school age children's responses for a number of reasons: (a) early elementary school age is above an age span when resting HRV is increasing (Finley and Nugent, 1995), (b) developmental comparisons with early elementary school age are also similar to past behavioral studies evaluating executive functioning age differences (Luciana and Nelson, 2002; Huizinga et al., 2006), and (c) the age of our sample is before the final adolescent growth spurt in executive functioning abilities that occurs during adolescence (De Luca and Leventer, 2008).

CLINICAL SIGNIFICANCE OF THE STUDY

The current study makes use of multiple tasks that are used for clinical assessment of executive functioning abilities. It may be helpful clinically to know which of these tap into the form of cognitive effort indexed by HF-HRV, and the neural circuitry that underlies the HF-HRV response. Specifically interesting would be if a cognitive process were to in past literature activate frontal regions, but not elicit a parametric change in HRV with difficulty. The spatial resolution of function MRI (fMRI) is such that it may be that the neural regions underlying HF-HRV are not utilized as may appear on fMRI studies, or it might show that that region is being used but not in the way that modulates HRV. These executive functions are important for a large number of clinical concerns, ranging from judging developmental delay, or deficit with a disorder such as ADHD, to assessing atypical aging, where executive functions may be early to decline. Certainly the current study will suggest if the measures used should be considered equivalent when administered clinically to children and adults. They may not if children's and adults' HF-HRV responds differently to increased executive demands.

HYPOTHESES

We hypothesize that, in all tasks, incremental increases in executive functioning load will result in both adults and children presenting incremental decreases in HRV power and behavioral performance. Children may be less able to manage increased executive loads because of their undeveloped frontal control and may, therefore, have smaller changes in quality of performance with increasing executive load. This underdeveloped frontal control may also lead to children's HRV being less controlled and efficient, with their responsitivity being less incrementally locked to increases in executive load. Whether children's HRV will be higher or lower than adults is exploratory.

MATERIALS AND METHODS

PARTICIPANTS

Data were analyzed from 25 children (16 male, 6–10 years, *M* age = 8.6 years) from local schools and 34 adults (19 male, 18–25 years, *M* age = 22.0 years) from introductory psychology courses. Child participants were recruited through flyers posted at graduate student on-campus housing, since this housing is often utilized by graduate students with children. Adult participants were recruited through an undergraduate psychology subject pool. Adult and child samples included participants who, according to a self/guardian report questionnaire, were in good present and past health and currently taking no medications. All participants were recruited and tested using procedures in accordance with the Ethical Guidelines of Psychologists and Code of Conduct of the American Psychological Association (1992) and approved by the university Institutional Review Board.

DESIGN AND PROCEDURE

Upon arrival at the university laboratory, adult participants or child–parent pairs heard a brief description of the study and underwent consent/assent procedures. Adult participants or child–parent pairs then answered a questionnaire about the participants' basic demographic data, current and past health, medical/psychiatric diagnoses, and medications the participants were currently taking. The experimental session (∼45 min) then began.

The experimental session consisted of the researcher briefing the participants about the tasks, referred to as "puzzle games," and the opportunity to earn a performance bonus of up to $5. This bonus was in addition to the standard compensation of $5 for children and class credit for adults. All participants were encouraged equally but were not informed about their progress toward performance bonuses until the end of the experiment. Encouragement and financial incentive were used to address potential decreased vigilance, engagement, and/or effort across the session, which is a major concern when testing child participants. Use of financial incentive was particularly crucial in the current study due to the past research finding that young school age boys' performance and HRV revealed more attention to task when the children were offered monetary reward (Suess et al., 1997).

The experimenter escorted the participants to a sound-attenuated booth and fitted the participants with electrocardiogram (ECG) electrodes and a respiration gage belt. Participants were instructed to refrain from speaking and making non-task-related movements during data-collection/task periods. Participants then began the three computerized executive function tasks, the Day/Night Stroop, the Tower of London, and the N-Back, with the order of the tasks determined by a Latin square, counterbalanced design. There was no significant evidence that child or adult participant groups performed more poorly on tasks later in the session[1]. Difficulty conditions within each task were completed in

[1]An Age × Task × Order ANOVA revealed no significant Order × Task or Age × Order × Task interactions on any of the performance measures ($ps > 0.05$).

order of increasing executive load levels to reduce discouragement, a concern especially in children.

Before each difficulty condition, there was an experimenter-participant interactive break during which the experimenter provided encouragement and instructions. Instructions using a standard script and pictures, either in a flip-book or on a computer screen, consisted of the experimenter explaining the stimuli, responses, and objectives for the next level of difficulty. Next, the experimenter demonstrated the difficulty condition and the participants were given two opportunities to practice this difficulty condition. The experimenter corrected and guided the participants if they performed incorrectly during the practice opportunities. If participants' responses on the practice periods revealed a lack of understanding of the condition, the experimenter repeated the instructions, example, and practice session. Only two children required re-instruction, both on the N-Back task.

EXECUTIVE FUNCTIONING TASKS

Our tasks were designed to: (a) be appropriate for both children and adults, (b) tap into the cognitive function of interest across a range of difficulty, and (c) be free of either periodic stimuli or large motor movements that could modify the participants' heart rate patterns. Tasks were identical for children and adults. In all tasks conditions were presented in periods of 3-min each so as to equate temporal conditions in evaluating the HRV.

Inhibition task (Day-Night Stroop)

The task employed was a variant of the standard color-word Stroop (1935) task, which is widely used to measure response inhibition in adults. Although letters are easily recognized by children as young as 6 years of age, reading automaticity is achieved later in development (Saint-Aubin et al., 2005). This lack of reading automaticity makes the color-word version of the Stroop task less valid for children. Although picture-based Stroop variants, including the Day/Night Stroop task we employed, are more common in developmental research, they are effective at eliciting difficulty in automatic response inhibition from adults as well as children. This is evident as slower response times in the stimulus–response conflict condition (Diamond and Taylor, 1996; Diamond and Kirkham, 2005; Davidson et al., 2006).

In the most common administration of the Day/Night Stroop (Gerstadt et al., 1994) participants speak either matching ("day" to a picture of day) or opposite ("day" to a picture of night) responses to simple, colorful drawings. Our computerized version of this task required only a mouse click to make their picture selection and this allowed for the recording of respiration and HRV without contamination from speech-related artifact and allowing us millisecond response accuracy; otherwise it was very similar to the spoken Day/Night Stroop. Participants were presented with a sequence of images, cartoons of either day or night, appearing one at a time in the upper portion of the computer screen. In the lower portion of the screen were two smaller images, one of day and one of night, which served as response buttons when left clicked. Participants used a computer mouse to click on the matching picture in the control difficulty condition and the opposite picture in the inhibition difficulty condition. Following the response, the next picture in the series of images appeared in 500–2000 ms, with the inter-stimulus-intervals independently randomized for each participant.

During the inter-stimulus interval, participants moved their mouse cursor to a bulls-eye image located between the two response images. This prevented anticipatory movements and held constant the movement distance for each response button. The matching (control) difficulty condition aways preceded the mismatching (response inhibition) difficulty condition, with each condition period being 3 min. The instructions for both difficulty conditions emphasized responding both quickly and correctly.

Planning task (Tower of London)

The Tower of London is a task used clinically and experimentally with both children and adults to measure multi-step planning (Shallice, 1982; Krikorian et al., 1994; Anderson et al., 1996; Berg and Byrd, 2002; Berg et al., 2006, 2010). The original task apparatus consisted of three balls, red, blue, and green, placed on three pegs which can hold one, two, or three balls, respectively. The task objective is to transform an initial ball arrangement to match a goal ball arrangement in as few single-ball movements as possible.

In our computerized version of the task, the initial arrangement appeared as a large image at the bottom of the screen, and the goal arrangement appeared as a small image at the top of the screen. The minimum number of moves necessary to reach the goal position appeared in a box on the far right of the screen. Participants could begin solving problems as soon as they appeared, though they were encouraged to solve the problem within the minimum number of moves but to continue working on a problem until it was solved, even when they made more than the minimum number of moves required. We chose this administration in order to allow for detailed examination of performance for planfulness (Berg et al., 2006).

To move each ball from peg to peg, the participants made a small hand movement, a drag and drop motion with the computer mouse. When the goal arrangement was reached, the participants clicked a button labeled "Done," which appeared in the top right corner of the screen. The computer program prevented participants from breaking the rules, placing balls off of pegs or placing too many balls on a peg. Participants were presented three increasing planning load difficulty conditions of the Tower of London – problems requiring a minimum of 4, 5, and 6 moves for an optimal solution. In each difficulty condition, participants continued to solve problems with no maximum of that difficulty level until the difficulty condition period of 3 min was complete. These problems were selected based on minimum number of moves required to solve most efficiently, which is a strong predictor of difficulty (Berg et al., 2010). Unfortunately these data were collected before problem selection became based on other problem parameters such as goal and end start position or subgoals (Kaller et al., 2004). These problem aspects not being controlled may have contributed noise to our difficulty levels.

Working memory task (N-Back)

In order to examine participants' responsitivity to increasing WM load, participants performed four increasingly difficult conditions

of the N-Back task. Various versions of this task are commonly used as a measure of WM updating both with adults (e.g., Gevins et al., 1997; McEvoy et al., 1998; Müller et al., 2002) and children (Nelson et al., 2000; Vuontela et al., 2003, 2009; Astley et al., 2009). It is feasible for children 6 years and older to perform this letter-based memory task due to their ability to recognize individual letters (Christian et al., 2000). In the current study, practice trials demonstrated that all child participants were able to recognize the stimulus letters.

In all difficulty conditions, participants viewed a light blue computer screen with a sequence of black upper and lower case letters appearing one at a time in the middle of the screen. There were 51 stimuli in each difficulty condition, one third of which were targets. The stimulus duration was 500 ms and the inter-stimulus-intervals varied from 300 to 1600 ms. Both inter-stimulus interval and target position were randomized independently for each participant. In all difficulty conditions, the participants were instructed that, following each stimulus, they were to press one of two computer keyboard keys: either a green key with their left index finger for a target or a red key with their right index finger for a non-target. Participants were instructed that they should respond to every single stimulus with their best answer, even if they were uncertain. Correctness, not speed of responses, was emphasized, though participants were told that non-responses would be considered incorrect.

Participants performed four difficulty conditions requiring incrementally more WM load: 0-, 1-, 2-, and 3-back difficulty conditions, in that order. The definition of a target stimulus differed by difficulty condition. In the 0-back difficulty condition a target was a single letter presented before the response stimuli for that difficulty condition. For all other difficulty conditions, the participants referred back to their memory of the prior stimuli in order to determine whether or not the current stimulus letter was a target. For these difficulty conditions, a target was always the matching letter, and the letter's case was to be ignored. The position to check for this match was the letter 1, 2, or 3 positions back in the sequence depending on difficulty condition being tested, 1-, 2-, or 3-back, respectively. Each difficulty condition period lasted 3 min.

HRV RECORDING AND MEASUREMENT CALCULATION

During each 3-min difficulty condition period, heart rate and respiration were recorded as six consecutive, 30-s epochs. This epoch duration was chosen because it was appropriate for examining frequencies of interest, brief enough to lessen concerns about heart rate non-stationarities (Berntson et al., 1997), and identical to that used by earlier studies of children's HRV (Suess et al., 1997; Porges et al., 2007).

Electrocardiography was recorded using three 1 cm Ag/AgCl electrodes filled with Microlyte electrolyte gel and secured to the cleaned and lightly abraded skin (Nu-Prep gel) via an adhesive electrode collar. Electrodes were placed in a modified type II arrangement, with two active leads, one on the left ankle and the other on the right collarbone, and a ground lead on the mastoid bone behind the left ear. Pilot testing determined that this placement allowed for unobtrusive electrode application as well as a clear EKG signal with sharply peaked R waves.

The EKG signal was amplified 1000x with a Coulbourn S75-01 bioamplifier, then band pass filtered from 8 to 40 Hz in order to minimize drift, movement artifact, and 60 Hz noise. A custom-designed peak detector was used to find the peak of the R-waves and transform the peak trigger to a short TTL pulse. R-R intervals (time between R-wave-triggered TTL pulses) were recorded with 1 ms accuracy by a custom program.

R-R interval timings were processed offline. All R-R interval editing and checking was conducted by trained and reliable research assistants who edited data unaware of the participants' task orders. A custom program was used to display the sequence of R-R intervals and edit artifactual intervals (dividing combined R-R intervals or combining R-R intervals interrupted by false triggering of the peak detector). Corrected data were re-checked for errors. Specific care was taken in the editing of R-R interval artifacts due to the large impact even a single artifactual R-R interval can have on the outcome of HRV calculations (Berntson and Stowell, 1998).

Six 30-s epochs were recorded for each difficulty condition of each task. Uneditable and/or unusable heart rate epochs were extremely rare. Three children and one adult had R-R recording artifacts that could not be clearly, reliably edited, resulting in one or more unusable 30-s epochs of data. Overall, 99.6% of the children's data and 99.8% of the adults' data were included in the analyses.

Using a custom BASIC program, the corrected series of R-R intervals during each 30-s epoch was re-sampled into 250-ms bins. This transformed the R-R intervals into a time-based sequence of R-R interval data, a series of densely sampled weighted R-R intervals for each 250 ms during the 30 s epoch. Further offline processing of R-R interval samples was conducted using Microsoft Excel. Linear trends were removed from each epoch's time-based sequence of R-R intervals using a linear regression model. Each epoch's de-trended time series was subjected to a fast Fourier transform (FFT) to obtain the power present in the different spectral bands. HRV values for each difficulty condition were calculated by taking the natural log of each 30-s epoch's absolute power in the frequency band associated with respiration, then averaging together these natural logs across the six 30-s epochs recorded during each difficulty condition.

Slightly different respiration frequencies were examined for child and adult groups, 0.15–1.03 Hz for adults (a frequency band common to adult studies of HRV and RSA; see Berntson et al., 1997) and 0.28–1.03 Hz for children (a frequency band similar to the frequency band 0.24–1.04 Hz common to child studies of HRV and RSA; see Hickey et al., 1995; Suess et al., 1997; Porges et al., 2007). These frequencies were empirically confirmed from respiration recordings taken during the current study[2].

[2]Respiration was recorded contemporaneously with the heart rate using a 10 cm mercury strain gage attached to an elastic belt wrapped snuggly around the participants' chests. The gage was attached to a Parks Electronics Model 270 plethysmograph transducer box to convert the signal to recordable voltages. Respiration data was sampled by a Tecmar 12 bit A/D converter at a rate of 10 Hz for the 30-s epoch and, using DOS-based custom software, recorded onto a computer hard disk. An FFT was used to determine the power of the frequencies present within this respiration signal.

BEHAVIORAL PERFORMANCE RECORDING AND MEASUREMENT CALCULATION

A total of six behavioral measures were analyzed, two behavioral measures for each of the three tasks, a measure of correctness and a measure of speed (as suggested by Berg and Byrd, 2002). The behavioral measures calculated to capture Day-Night Stroop task performance were the proportion of correct responses given and the response time for all responses. The performance measures for the Tower of London were the number of perfect solutions (solved in the minimum moves possible) and the time taken to solve Tower of London problems (from first move to last). The N-Back performance measures were the proportion of correct responses and response time for all responses.

RESULTS

PRELIMINARY DATA PROCESSING

Raw behavioral performance variables were analyzed for strong skew. Variables where the absolute value of the mean of the Fisher kurtosis score divided by the standard error was two or larger ($z$skew = | Skew|/SEskew) were transformed using the natural log (Royston, 1992). All behavioral scores, excluding the Stroop proportion correct responses and Tower of London number perfect solutions, were transformed.

The few missing scores, 2% of the data, resulted from random, non subject-specific causes (e.g., computer error during testing, uneditable data, corrupted computer file). In order to maintain the sample size across tasks, missing scores were estimated using the expectation–maximization (EM) method (Dempster et al., 1977; Little and Rubin, 1987).

All measures, including correctness, response speed, and HRV, were then converted into "responsitivity scores" to test directly our hypotheses about the developmental responsitivity to increased executive load in performance and HRV. For the Day/Night Stroop task, responsitivity was calculated as the inhibition (opposite) difficulty condition relative to the control (matching) difficulty condition. For the Tower of London task, responsitivity in the 5-move and 6-move difficulty conditions was examined relative to the 4-move difficulty condition. For the N-Back task, responsitivity in the 1-, 2-, and 3-back difficulty conditions were examined relative to the 0-back difficulty condition. All raw scores and transformed responsitivity scores are presented in **Table 1**.

All scores were evaluated for possible ceiling and floor effects as reported below and noted in **Table 1** when significant. Gender differences were examined separately for children and adults using between-subject two-tailed t-tests. Only two significant gender differences were found: adult females were more reactive in their n-back, 1-back condition proportion correct (solving a smaller proportion correct relative to 0-back than males) and child females were more reactive in their n-back, 2-back condition HRV (more suppression of HRV). Genders were combined for all analyses except these two measures, which were also analyzed for potential gender interactions.

In order to determine if there were significant age differences within our dependent variables, we conducted a median split of our child group based on age [$N = 12$ younger: $M(SE) = 7.69$ (0.23), $N = 13$ older $M(SE) = 9.61(0.12)$]. Two of these one-tailed t-tests were significantly different. Younger children differed in responsitivity for 2-back and 3-back conditions of the N-back [$ts(23) > 2.05$, $p < 0.029$]. Younger children also showed fewer perfect solutions on the Tower of London than older children [$t(23) = 2.04$, $p = 0.024$]. For these measures a solutions Age Group × Condition analyses will be conducted.

All analyses with repeated measures were Greenhouse–Geisser corrected. When age differences were *a priori* hypothesized, analyses were conducted using single tail t-tests.

AGE AND LOAD DIFFERENCES ON PERFORMANCE AND HRV RESPONSITIVITY

Inhibition of automatic response (Day-Night Stroop task)

This task had one responsitivity level due to the task design of a single condition of increased inhibition load (mismatching condition) compared to the condition of no inhibition load (matching condition). For the dependent variable of Stroop proportion of correct responses, both children and adults performed at ceiling, above 0.97 correct responses. Further analysis of this ceiling-level measure was not performed.

A 1-way (Age) analysis of variance (ANOVA) was conducted for the Stroop dependent variable of reaction time. For reaction, time children were more reactive than adults in their slowing of responses to the inhibition condition [$F(1,59) = 11.50$, $p = 0.001$, partial $\eta^2 = 0.168$]. For HRV, means suggested the adults' responses were somewhat more suppressed during the inhibition condition than those of the children, but this age difference was not significant [$F(1,59) = 1.51$, $p = 0.225$, partial $\eta^2 = 0.026$]. As indicated in **Table 1**, children's, but not adults', HRV responsitivity in the inhibition condition was significantly suppressed relative to the control condition [$t(24) = 1.35$, $p = 0.042$].

Multistep planning (Tower of London task)

This task had two levels of responsitivity (two levels of increased difficulty) due to the task design of 5- and 6-move conditions each being compared relative to the easiest, 4-move condition. Children and adults were compared in their responsitivity of the number of perfect solutions with increased planning load using an Age × Difficulty Condition (2 × 2) ANOVA. Though the means suggested that adults had a higher number of perfect solutions, the main effect for age was not significant [$F(1,57) = 1.94$, $p = 0.169$, partial $\eta^2 = 0.033$]. Across age groups, the number of perfect solutions decreased from 5-move to 6-move problems [main effect planning load: $F(1,57) = 17.90$, $p < 0.001$, partial $\eta^2 = 0.239$], but this decrease did not differ between children and adults [Age × Planning Load interaction: $F < 1$].

Since analyses presented above indicated a significant age difference within the child group for this measure of 6-move number of perfect solutions, each child subgroup is separately compared to the adult group. The younger child group was significantly different in suppression in number perfect solutions compared to adults [$t(16) = 2.08$, $p = 0.026$]. The older children were not significantly different in suppression of number of perfect solutions compared to adults [$t < 1$].

Children and adults were compared in their responsitivity of slowing of solution time with increased planning load using an Age × Difficulty Condition (2 × 2) ANOVA. A significant main effect of age revealed that adults' speed of solution was more

Table 1 | Descriptive statistics for raw and reactivity/responsitivity performance and HRV measures for each task.

			Raw scores				Skew-corrected reactivity scores			
			Children		Adults		Children		Adults	
Task	**Dependent variable**	**Difficulty**	***M***	***SE***	***M***	***SE***	***M***	***SE***	***M***	***SE***
Stroop	Proportion correct	Control	1	0.001	0.99	0.003	–	–	–	–
		Inhibition	0.99[c]	0.002	0.97[c]	0.006	−9.458[a]	1.080	−5.500[a]	0.662
	Response time (ms)	Control	929.188	32.702	590.305	10.779	–	–	–	–
		Inhibition	1216.541	59.675	703.421	16.89	0.114	0.011	0.075	0.005
	HRV	Control	7.844	0.109	7.781	0.068	–	–	–	–
		Inhibition	7.798	0.115	7.674	0.07	−0.046[b]	0.034	−0.107	0.034
Tower of London	Num perfect solutions	4-Move	4.826	0.513	8.411	0.544	–	–	–	–
		5-Move	2.522	0.301	4.382	0.437	−0.177[a]	0.060	−0.344[a]	0.053
		6-Move	0.957	0.204	2.5	0.373	−0.452[a]	0.058	−0.573[a]	0.050
	Time to solve (sec)	4-Move	23.42	3.137	17.557	0.981	–	–	–	–
		5-Move	43.631	8.033	38.396	3.153	0.223	0.07	0.314	0.03
		6-Move	41.969	5.848	66.503	7.574	0.239	0.061	0.508	0.046
	HRV	4-Move	7.901	0.117	7.712	0.064	–	–	–	–
		5-Move	7.889	0.126	7.751	0.062	−0.012[b]	0.036	0.039[b]	0.026
		6-Move	7.872	0.134	7.722	0.063	−0.029[b]	0.043	0.010[b]	0.035
N-Back	Proportion correct	0-Back	0.861	0.022	0.976	0.005	–	–	–	–
		1-Back	0.78	0.024	0.944	0.007	−0.016	−0.005	−0.007	0.001
		2-Back	0.666	0.026	0.915	0.011	−0.038	0.005	−0.013	0.002
		3-Back	0.63	0.017	0.819	0.014	−0.038	0.004	−0.024	0.003
	Response time (ms)	0-Back	627.146	26.899	546.938	23.519	–	–	–	–
		1-Back	729.108	46.76	633.063	26.405	0.055	0.014	0.064	0.009
		2-Back	757.084	53.034	798.115	40.769	0.066	0.023	0.157	0.017
		3-Back	684.001	45.601	866.159	40.767	0.025	0.021	0.194	0.016
	HRV	0-Back	8.099	0.116	7.768	0.074	–	–	–	–
		1-Back	8.034	0.123	7.791	0.071	−0.064	0.052	0.023	0.035
		2-Back	7.984	0.113	7.688	0.062	−0.114	0.05	−0.08	0.044
		3-Back	8.053	0.117	7.609	0.062	−0.046	0.054	−0.159	0.041

[a] *Reaction score not skew-corrected since not necessary.*
[b] *Value not significantly different than baseline value of 0 (single sample t-tests conducted for each age group, ps < 0.05 relative to baseline).*
[c] *Value not significantly different than ceiling value of 1.0 (single sample t-tests conducted for each age group, ps < 0.05 relative to maximum score).*

reactively slowed than children's [$F(1,57) = 7.95$, $p = 0.007$, partial $\eta^2 = 0.122$]. A significant main effect of difficulty revealed that participants' responsitivity in speed of solution was slower with increased planning load [$F(1,57) = 9.39$, $p = 0.003$, partial $\eta^2 = 0.141$]. Children and adults differed in their solution time responsitivity with increased planning load [Age × Difficulty Condition interaction: $F(1,57) = 6.71$, $p = 0.012$, partial $\eta^2 = 0.105$]. When age groups were tested separately, adults exhibited significantly slowed solution time from 5-move to 6-move problems [$t(33) = 5.19$, $p < 0.001$], but children did not [$t < 1$]. See **Figure 1**.

The Age × Difficulty Condition (2 × 2) ANOVA examining HRV found that HRV was not significantly reactive in its suppression with increased planning load [main effect of difficulty condition: $F(1,57) = 1.37$, $p = 0.247$, partial $\eta^2 = 0.023$]. As planning load increased, neither children nor adults changed their HRV suppression [main effect of age: $F < 1$], nor did age group and difficulty condition significantly interact on this measure [Age × Difficulty Condition interaction: $F < 1$]. When age groups were tested independently, neither adults' nor children's HRV values were significantly different than baseline [adults: $ts(33) < 1.37$, $ps > 0.179$; children: $ts(24) < 1$].

Working memory (N-Back task)

Floor effects for the proportion correct raw scores were assessed by comparing results to the chance performance level of 0.50, the result if target or non-target buttons were randomly pressed. For both age groups, performance in all conditions was significantly better than chance ($ps < 0.001$). Though adults' proportion

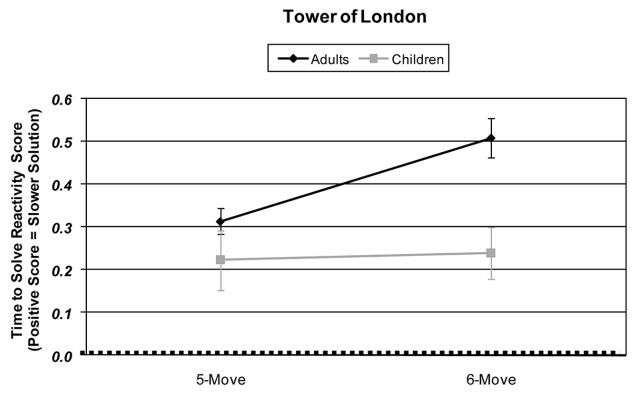


FIGURE 1 | Age and planning-load-level differences in Tower of London skew-corrected behavioral performance responsitivity as measured by time taken to solve problems (sec). Responsitivity for each difficulty level is calculated by subtracting that difficulty condition's performance from the 4-move difficulty condition performance.

correct was high for the 0- and 1-back conditions, their proportions correct were significantly below the ceiling value of 1.00 [$ts(33) > 4.56$, $ps < 0.001$].

Children's and adults' proportion correct responsitivity was examined across 1-, 2-, and 3-back conditions using an Age × Difficulty Condition (2 × 3) ANOVA. Children's decrease in performance was larger than adults [main effect of age: $F(1,57) = 21.41$, $p < 0.001$, partial $\eta^2 = 0.273$]. For both age groups, there was a decrease in proportion correct with increasing WM load [main effect of difficulty condition: $F(2,114) = 30.41$, $p < 0.001$, partial $\eta^2 = 0.348$]. The pattern of proportion correct responsitivity differed for children and adults [Age × Difficulty condition interaction: $F(2,114) = 5.40$, $p = 0.008$, partial $\eta^2 = 0.081$]. Adults' proportion correct decreased with each level of difficulty [$ts(33) > 2.38$, $ps < 0.03$]. Children's proportion correct decreased from 1- to 2-back and 1- to 3-back [$ts(24) > 4.38$, $ps < 0.001$], but did not differ from 2- to 3-back [$t < 1$]. The children's within-group variability was larger than that of the adults.

Children's and adults' response time responsitivity was examined across 1-, 2-, and 3-back conditions using an Age × Difficulty Condition (2 × 3) ANOVA. Adults' responsitivity was more slowed than children's [main effect of age: $F(1,57) = 18.65$, $p < 0.001$, partial $\eta^2 = 0.247$] and response time responsitivity differed among difficulty conditions [main effect of difficulty condition: $F(2,114) = 14.96$, $p < 0.001$, partial $\eta 2 = 0.208$]. For the participants as a whole' response speed responsitivity differed in response to increased WM difficulty conditions [Age × Difficulty Condition interaction: $F(2,114) = 28.24$, $p < 0.001$, partial $\eta^2 = 0.311$]. Adults' response time responsitivity slowed with each increase in WM difficulty, all pair-wise comparisons were significant [$ts(33) > 2.68$, $p < 0.011$], and each of these adult responsitivity scores was significantly slowed relative to control [$ts(33) > 7.19$, $ps < 0.001$].

Children's responsitivity in reaction time did not slow from 1- to 2-back [$t(24) < 1$], and their reaction time responsitivity was actually significantly *faster* in 3-back than 2-back [$t(24) = 3.19$, $p = 0.004$]. This pattern resulted in the 3-back reaction time responsitivity nearing significance in its difference from 1-back condition responsitivity [$t(24) = 2.01$, $p = 0.055$]. Children's response time responsitivity was slower than baseline in the 1- and 2- back conditions [$ts(24) > 2.81$, $ps < 0.011$], but not so in the 3-back condition [$t(24) = 1.16$, $p = 0.128$]. There was generally larger within-group variability in the child data.

An Age × Difficulty Condition (2 × 3) ANOVA of HRV responsitivity revealed no significant main effects of age [$F(1,57) < 1$], but did reveal a main effect of difficulty [$F(2,114) = 4.57$, $p = 0.014$, partial $\eta^2 = 0.074$]. Additionally, adults and children differed in their HRV responsitivity to increasing WM difficulty resulting in a significant age by difficulty interaction [$F(2,114) = 5.86$, $p = 0.005$, partial $\eta^2 = 0.093$]. Adults' HRV responsitivity was suppressed with each increasing level of WM difficulty [1- vs. 2-back: $t(33) = 2.96$, $p = 0.006$; 2- vs. 3-back: $t(33) = 2.54$, $p = 0.016$; 1- vs. 3-back: $t(33) = 4.69$, $p < 0.001$]. Although the children's means in the 1- and 2-back conditions appeared reactive to WM load, the children's HRV responsitivity did not significantly differ among WM difficulty levels (all $ts < 1.46$, $ps > 0.158$). The adults' and children's 1-back HRV responsitivity was not significantly different from baseline, and, interestingly, the children's 3-back HRV responsitivities were not significantly suppressed below baseline [$ts < 1.24$, $ps > 0.113$]. See **Figure 2**.

Because the younger and older children in earlier analyses showed different performance on the 2- and 3-back conditions, each age subgroup was compared to adults. The younger subgroup of children differed in significance compared to the response time of adults for both 2- and 3-back conditions [$ts(45) > 4.12$, $ps < 0.001$]. The older children were not significantly different than adults in suppression of response time for the 2-move condition [$t(44) = 1.25$, $p = 0.219$], but these older children differed in their performance time on the 3-back problems [$t(44) = 3.84$, $p < 0.001$].

DISCUSSION

The current study builds on past literature by examining developmental differences in HRV responsitivity to increased executive load. Both child and adult groups were assessed across multiple executive function tasks focused on three critical facets of executive functioning. Each task was designed to incrementally increase the executive control necessary for correct and rapid responses, and also assess those executive functions found to be related to executive effort and HRV in the past literature (Thayer et al., 2009). Our task designs were validated by behavioral results. For both age groups and all executive function tasks, behavioral performance was suppressed with increased executive load. Generally, adults were more behaviorally reactive, showing larger decreases in speed of performance with increasing load, as compared to children.

Results for the HRV responding were more complex. The two tasks that required a series of discrete, timed response in relatively rapid succession – the inhibition and WM tasks – produced HRV suppression that was reactive to increased executive cognitive load. It may be that the time pressure of fast responding caused a cognitive state requiring an overwhelming amount of executive control. When HRV suppression was produced it was again more reactive

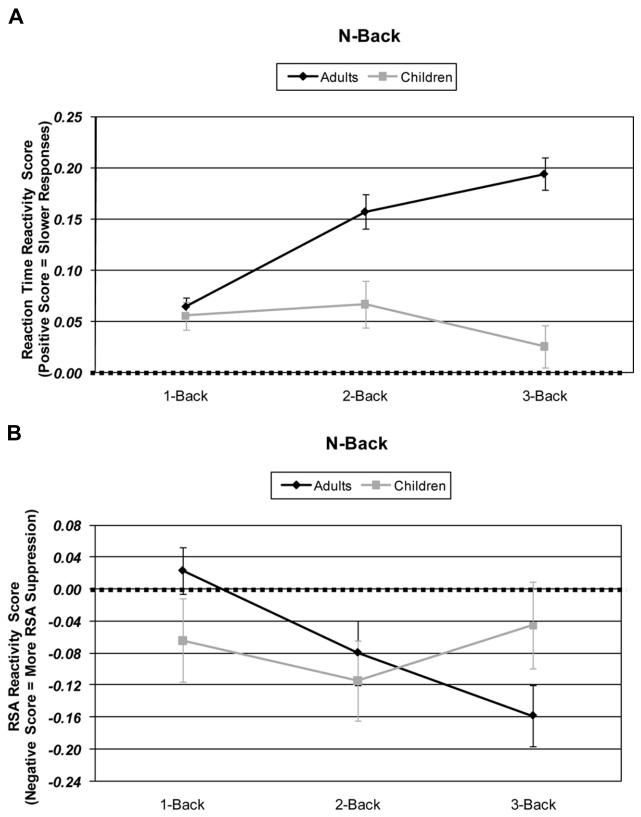


FIGURE 2 | Age and working memory-load-level differences in: (A) N-Back behavioral performance responsitivity and as measured by response time (ms). (B) N-Back physiological response reactivity as measured by high frequency heart rate variability (HRV). Skew-corrected reactivity for each difficulty level is calculated by subtracting that difficulty condition's HRV from the 0-back difficulty condition HRV.

overall and also more reactive to increased load in adults [similar to reduced HRV responsitivity in children during a version of the Stroop task by Mathewson et al. (2010)]. The more complex multistep planning task, which required slow, self-paced responses over a longer time than the other tasks, showed behavioral responsitivity while not producing any significant HRV responsitivity for either age group.

The N-back, with its multiple levels of difficulty across a wide range of WM loads, may offer the most insight into developmental differences in HRV responsitivity to executive load. With this task, it is possible to examine multiple levels of *effective difficulty*, which can also be conceptualized as age-group-specific levels of moderate and high difficulty [an analysis approach suggested by Katsoni et al. (2006)]. The 1- and 2-back conditions can be reasonably viewed as moderate and high effective difficulty levels in children, and 2- and 3-back conditions can be reasonably viewed as moderate and high difficulty levels in adults. With this assignment, a different comparison across age can be assessed. When this age-specific difficulty adjustment is made, similarity rather than difference appears (see **Figure 3**). Specifically, patterns of HRV suppression are similar between age groups. This suggests that when subjective difficulty requires similar amounts of effort, children and adults may show similar effort-related HRV suppression.

Of course, the obvious question with this interpretation is: "What about the 3-back with children – isn't it also very difficult?" The reason we exclude this condition here is that we interpret the whole of the results, behavioral as well as HRV, as an indication that the children appeared to be overwhelmed by the most demanding, 3-back condition of the N-back task. The strongest evidence of this was that behavioral performance was near chance. The children may have given up mental effort during this most difficult condition. The result to be expected, if this is the case, is little HRV suppression, just what we found.

HRV AS AN INDEX OF EXECUTIVE EFFORT

Adults' HRV responsitivity increased with increased executive loads in the inhibition and WM tasks, but not the planning task. These patterns suggest that HRV does index some forms of executive effort, perhaps those that require assessing a rapid series of discrete stimuli while processing and responding in a speeded manner with a relatively high density of responses, similar to those tasks used in past studies of HRV-Executive Function relationships (see Thayer et al., 2009 for a review). Speeded and high density responses were characteristics of our inhibition and WM tasks. Slower, self paced, and multi step responses required by our planning task may require a form of executive functioning not indexed by HRV. This implies that HRV suppression is sensitive to a specific form of attentional control requiring vigilance to a rapidly change course of stimuli not under the participant's control rather than a largely stationary stimulus where responding is under the participant's control. An alternative administration of a planning task with more rapid presentation of problems and a single button response would be more similar to our inhibition and WM tasks and would allow us to determine further if planning is an executive function reflected in HRV responsitivity. We

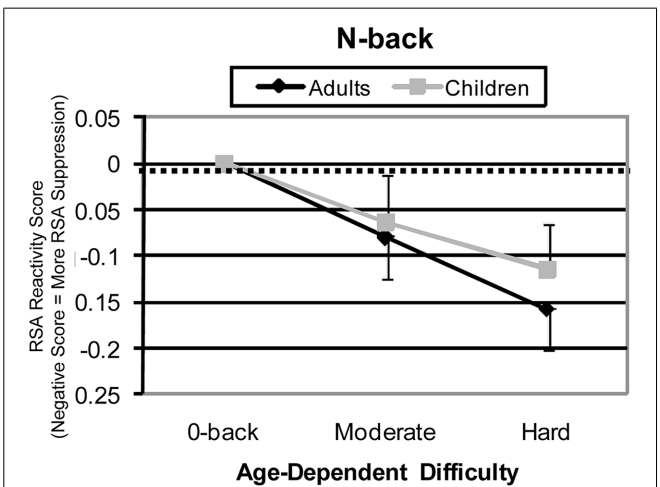


FIGURE 3 | N-Back high frequency HRV reactivity for effective (age-equated) difficulty levels. For adults, medium, and hard difficulty conditions were 2-back and 3-back. For children, medium and hard difficulty conditions were 1-back and 2-back conditions. Adjusted patterns are shown for HRV reactivity.

also could have offered a simpler planning baseline, such as 1- and 2-move problems, and then perhaps we would have seen planning difficulty differences. Finally, this data was collected before Berg et al. (2010) as well as Kaller et al. (2004) published other problem parameters other than minimum number of moves that determine difficulty. Not controlling for these other parameters may have created noise and overlap between out TOL conditions preventing a clear parametric design for this task.

The large amount of children's HRV variance during the planning task may have resulted from the multiple slow responses for a single solution and from variance in the strategy/approach to the task. For example, the current study's instructions and reward schedule encouraged planfulness, but the Tower of London task, like other tower-transfer planning tasks, can be approached with strategies requiring more or less multistep planning. Participants may use lower planning effort strategies that still reach the goal using strategies based on surface appearance and making random moves hoping to "chance upon" the solution (Berg et al., 2006). During the most difficult planning task conditions, child participants may have been switching among approaches requiring more and less executive effort, with some moves or sequences of moves during the solution period being more planful than others. There is some evidence in the data to support that children were varying more greatly in switching among different, more and less effortful approaches or strategies when faced with the most difficult planning load. This variability was larger in the most difficult Tower of London condition (0.134) than in the most difficult conditions of the Stroop (0.115) and N-back (0.117) tasks. This pattern of variances was not present in the adults (0.070, 0.063, and 0.062 relatively).

This interpretation of HRV's sensitivity to strategy also matches well with the pattern of behavioral and HRV responsitivity that children displayed during the most difficult condition in the WM task. When overwhelmed with the most difficult, 3-back condition of the WM task, the children appear to have switched to a less executive/effortful strategy for this task, perhaps responding based on familiarity rather than encoding each item (Speer et al., 2003).

ADDRESSING HYPOTHESES

We hypothesized that incremental increases in executive load would result in incremental decreases in behavioral performances and HRV. This pattern was present in the adults during the WM task, showing incrementally more suppressed HRV along with incrementally poorer performance. This incremental HRV change may be most evident in the task that had many (4) levels of difficulty and which required vigilance and speeded responding to rapidly presented stimuli. Except for the most difficult condition, where children were overwhelmed, children's responses were also incremental in appearance.

CLINICAL SIGNIFICANCE OF FINDINGS

There are some clinical ramifications of the current study, specifically when clinicians are determining test design to monitor what executive functions may be at deficit. Those that use a more time pressure, speeded response may tax a different form of executive functioning than a task that is self-paced. Developmentally, this study underscores the importance of choosing age-appropriate difficulty levels of executive functioning tasks, as the giving-up behavior in the most difficult N-back condition, poor performance can occur not because the participant is trying and struggling, but simply because they are giving up.

LIMITATIONS OF CURRENT RESEARCH AND FUTURE DIRECTIONS

The most serious limitations for this study come from the planning task, where there were no high frequency phasic HRV differences found with increasing executive load. We hesitate to think that planning as an executive function is not indexed by HRV, but think that the way that we administered the planning task may have limited the HRV responsitivity. One potential design aspect that could have hidden HRV responsitivity is that difficulty levels were not spread far enough among easy, medium, and difficult conditions. Future studies may wish to vary planning difficulty as widely as WM difficulty, with baselines of 1 or 2 move problems, and difficulties ranged widely as it was between 0-back and 3-back. With this change in design we could compare very low planning load, moderate planning load, and high planning load. This may show one of the limitations of the parametric design, that a full range of difficulty must be presented.

Additionally, the pattern in the results where we saw HRV responsitivity in WM and inhibition executive functioning tasks, but not the planning task may have also revealed that high frequency phasic HRV is most sensitive to increases in executive function load when there is some time pressure in response, as there was in our Day-Night Stroop and N-back tasks. Perhaps we would have seen a planning difference if we had told the participants to solve as quickly as possible, or perhaps if we had given them a different variation on the Tower of London, one more similar to how it is used in fMRI studies where participants see the start and goal positions, solve problems covertly, in their mind's eye, and then respond either with a button press of how many moves it take or solving the problem with mouse movements (Unterrainer et al., 2004b).

This idea of speeded responding being more strongly indexed by HRV may relate to one of the other applications of HRV, to emotional regulation (Thayer and Lane, 2000) and specifically to anxiety (Appelhans and Luecken, 2008). It may be that the Stroop and N-back with their speeded responding were more anxiety provoking, than the planful moves approach that was the best approach for the TOL. The more difficult Stroop and N-back conditions may have caused more anxiety or emotional dysregulation than easier conditions, while with the TOL solving fewer moves did not cause less emotional dysregulation than more difficult conditions. This again points to future studies putting executive functions on an even field as to speeded response, with TOL having to be solved in the head as quickly as possible.

Our use of performance based reward, which participants did not see until the end of the task, may have also played a role in which executive tasks showed HRV responsitivity. In past literature, reward have been seen to make a difference in the performance of certain, gambling-related executive functioning tasks, that is for reward for a different odds-based game of chance context with preschoolers (Kerr and Zelazo, 2004). Concerns of the

reviewers suggest that future studies should be conducted to determine if reward, such as offered in our study change the anxiety level in the certain, time pressured tasks.

CONCLUSION

In sum, high frequency phasic HRV appears sensitive to increasing executive demand in adults and children for WM and inhibition tasks. The exception to this was in the WM condition that was too difficult for the children, where there performance reverted to chance levels, suggesting the children were just guessing responses, and their HRV returned closer to baseline. We were most surprised by the findings with the planning task, where there was no HRV responsitivity with increased planning load. We discussed above why that may be so, and how future studies can investigate if planning is truly an executive function that does not have an impact on HRV or if HRV is sensitive to some of the task parameters that a multi-step planning task may have, as compared to a simple, single button/single click time-pressured task, such as our N-back and Day-Night Stroop tasks.

The children's HRV was less reactive than adults suggesting that decreased frontal lobe involvement in these children may impact the sympathetic and parasympathetic systems such that there is decreased HRV responsitivity. This is somewhat surprising, as children's time locked evoked heart rate responses are larger than adults, children's HRV could have been more reactive (Byrd and Berg, 2002).

REFERENCES

Althaus, M., Mulder, L. J. M., Mulder, G., Aarnoudse, C. C., and Minderaa, R. B. (1999). Cardiac adaptivity to attention demanding tasks in children with a pervasive developmental disorder not otherwise specified. *Biol. Psychiatry* 46, 799–809. doi: 10.1016/S0006-3223(98)00374-6

Althaus, M., Van Roon, A. M., Mulder, L. J., Mulder, G., Aarnoudse, C. C., and Minderaa, R. B. (2004). Autonomic response patterns observed during the performance of an attention-demanding task in two groups of children with autistic-type difficulties in social adjustment. *Psychophysiology* 41, 893–904. doi: 10.1111/j.1469-8986.2004.00252.x

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *Am. Psychol.* 47, 1597–1611. doi: 10.1037/0003-066X.47.12.1597

Anderson, P., Anderson, V., and Lajoie, G. (1996). The Tower of London test: validation and standardization for pediatric populations. *Clin. Neuropsychol.* 10, 54–65. doi: 10.1080/13854049608406663

Appelhans, B. M., and Luecken, L. J. (2008). Heart rate variability and pain: associations of two interrelated homeostatic processes. *Biol. Psychol.* 77, 174–182. doi: 10.1016/j.biopsycho.2007.10.004

Astley, S., Richards, T., Aylward, E., Olson, H., Kerns, K., Brooks, A., et al. (2009). Magnetic resonance spectroscopy outcomes from a comprehensive magnetic resonance study of children with fetal alcohol spectrum disorders. *Magn. Reson. Imaging* 27, 760–778. doi: 10.1016/j.mri.2009.01.003

Baddeley, A. (1996). Exploring the central executive. *Q. J. Exp. Psychol.* 49A, 5–28. doi: 10.1080/713755608

Banich, M. T., Milham, M. P., Atchley, R., Cohen, N. J., Webb, A., Wszalek, T., et al. (2000). fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *J. Cogn. Neurosci.* 12, 988–1000. doi: 10.1162/08989290051137521

Beauchaine, T. (2001). Vagal tone, development, and Gray's motivational theory: toward an integrated model of autonomic nervous system functioning in psychopathology. *Dev. Psychopathol.* 13, 183–214. doi: 10.1017/S0954579401002012

Berg, W. K., and Byrd, D. L. (2002). The Tower of London spatial problem-solving task: enhancing clinical and research implementation. *J. Clin. Exp. Neuropsychol.* 24, 586–604. doi: 10.1076/jcen.24.5.586.1006

Berg, W. K., Byrd, D. L., McNamara, J. P., and Case, K. (2010). Deconstructing the tower: parameters and predictors of problem difficulty on the Tower of London task. *Brain Cogn.* 72, 472–482. doi: 10.1016/j.bandc.2010.01.002

Berg, W. K., Byrd, D. L., McNamara, J. P., and MacDonald, C. (2006). "New indices of planning abilities using the Tower of London," in *The Quantified Process Approach to Neuropsychological Assessment,* ed. A. Poreh (New York: Taylor & Francis).

Berg, W. K., DeLucca, T. L., Case, K., Byrd, D. L., and McNamara, J. P. H. (in press). "The impact of childhood development and adult aging on Tower of London performance," in *Edouard Lucas and his Tower of Hanoi: History, Mathematics and Cognition,* eds A. Danek, H. M. Hinz, and F. Gauzy (New York: Springer).

Berntson, G. G., Bigger, J. T. Jr., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., et al. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology* 34, 623–648. doi: 10.1111/j.1469-8986.1997.tb02140.x

Berntson, G. G., and Stowell, J. (1998). ECG artifacts and heart period variability: don't miss a beat! *Psychophysiology* 35, 127–132. doi: 10.1111/1469-8986.3510127

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi: 10.1037/0033-295X.108.3.624

Byrd, D. L., and Berg, W. K. (2002). Developmental differences in the preparatory heart rate response: childhood through adulthood. *Biol. Psychol.* 61, 271–276. doi: 10.1016/S0301-0511(02)00062-5

Case, R. (1992). The role of the frontal lobes in the regulation of cognitive development. *Brain Cogn.* 20, 51–73. doi: 10.1016/0278-2626(92)90061-P

Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Cohen, J. D., et al. (1997). A pediatric functional MRI study of prefrontal activation during performance of a Go- No-Go task. *J. Cogn. Neurosci.* 9, 835–847. doi: 10.1162/jocn.1997.9.6.835

Christian, K., Morrison, F. J., Frazier, J. A., and Massetti, G. (2000). Specificity in the nature and timing of cognitive growth in kindergarten and first grade. *J. Cogn. Dev.* 1, 429–448. doi: 10.1207/S15327647JCD0104_04

Critchley, H. D., Mathias, C. J., Josephs, O., O'Doherty, J. O., Zanini, S., Dewar, B.-K., et al. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain* 126, 2139–2152. doi: 10.1093/brain/awg216

Davidson, M. C., Amso, D., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006

De Luca, C. R., and Leventer, R. J. (2008). "Developmental trajectories of executive functions across the lifespan," in *Executive Functions and the Frontal Lobes: A Lifespan Perspective*, eds P. Anderson, V. Anderson, and R. Jacobs (Washington, DC: Taylor & Francis), 23–56.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* 39, 1–38.

Denot-Ledunois, S., Vardon, G., Perruchet, P., and Gallego, J. (1998). The effect of attentional load on the breathing pattern in children. *Int. J. Psychophysiol.* 29, 13–21. doi: 10.1016/S0167-8760(97)00086-X

Diamond, A., and Kirkham, N. (2005). Not quite as grown up as we would think. *Psychol. Sci.* 16, 291–297. doi: 10.1111/j.0956-7976.2005.01530.x

Diamond, A., and Taylor, C. (1996). Development of an aspect of executive control: development of the abilities to remember what I said and to "do as I say, not as I do."*Dev. Psychobiol.* 29, 315–334. doi: 10.1002/(SICI)1098-2302(199605)29:4<315::AID-DEV2>3.0.CO;2-T

Durston, S., Hulshoff Pol, H. E., Casey, B. J., Giedd, J. N., Buitelaar, J. K., and van Engeland, H. (2001). Anatomical MRI of the developing human brain: what have we learned? *J. Am. Acad. Child Adolesc. Psychiatry* 40, 1012–1020. doi: 10.1097/00004583-200109000-00009

Finley, J. P., and Nugent, S. T. (1995). Heart rate variability in infants, children, and young adults. *J. Auton. Nerv. Syst.* 51, 103–108. doi: 10.1016/0165-1838(94)00117-3

Fuster, J. M. (2002). Frontal lobe and cognitive development. *J. Neurocytol.* 31, 373–385. doi: 10.1023/A:1024190429920

Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action: performance of children 3–7 years old on a Stroop-like day-night test. *Cognition* 53, 129–153. doi: 10.1016/0010-0277(94)90068-X

Gevins, A., Smith, M. E., McEvoy, L., and Yu, D. (1997). High-resolution EEG mapping cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cereb. Cortex* 7, 374–385. doi: 10.1093/cercor/7.4.374

Gianaros, P. J., van der Veen, F. M., and Jennings, J. R. (2004). Regional cerebral blood flow correlates with heart period and high-frequency heart period variability during working memory tasks: implications for the cortical and subcortical regulation of cardiac autonomic activity. *Psychophysiology* 41, 521–530. doi: 10.1111/1469-8986.2004.00179.x

Hajcak, G., McDonald, N., and Simons, R. F. (2003). To err is autonomic: error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology* 40, 895–903. doi: 10.1111/1469-8986.00107

Hanlon, H. W., Thatcher, R. W., and Cline, M. J. (1999). Gender differences in the development of EEG coherence in normal children. *Dev. Neuropsychol.* 16, 479–506. doi: 10.1207/S15326942DN1603_27

Hansen, A. L., Johnsen, B. H., and Thayer, J. F. (2003). Vagal influence on working memory and attention. *Int. J. Psychophysiol.* 48, 263–274. doi: 10.1016/S0167-8760(03)00073-4

Hickey, J. E., Suess, P. E., Newlin, D. B., Spurgeon, L., and Porges, S. W. (1995). Vagal tone regulation during sustained attention in boys exposed to opiates in utero. *Addict. Behav.* 2, 43–59. doi: 10.1016/0306-4603(94)00044-Y

Huizinga, M., Dolan, C. V., and van den Molen, M. W. (2006). Age related change in executive function: developmental trends and latent variable analysis. *Neuropsychologica* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010

Jansma, J. M., Ramsey, N. F., Coppola, R., and Kahn, R. S. (2000). Specific versus nonspecific brain activity in a parametric N-back task. *Neuroimage* 12, 688–697. doi: 10.1006/nimg.2000.0645

Jennings, J. R., and Yovetich, N. (1991). "The aging information processing system: autonomic nervous system indices," in *Handbook of the Psychophysiology of Information Processing: An Integration of Central and Autonomic Nervous System Approaches*, eds J. R. Jennings and M. G. H. Coles (Chichester: John Wiley & Sons), 706–727.

Johnsen, B. H., Thayer, J. F., Laberg, J. C., Wormnes, B., Raadal, M., Skaret, E., et al. (2003). Attentional and physiological characteristics of patients with dental anxiety. *J. Anxiety Disord.* 17, 75–87. doi: 10.1016/S0887-6185(02)00178-0

Kaller, C. P., Rahm, B., Spreer, J., Mader, I., and Unterrainer, J. M. (2008). Thinking around the corner: the development of planning abilities. *Brain Cogn.* 67, 360–370. doi: 10.1016/j.bandc.2008.02.003

Kaller, C. P., Rahm, B., Spreer, J., Weiller, C., and Unterrainer, J. M. (2011). Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. *Cereb. Cortex* 21, 307–317. doi: 10.1093/cercor/bhq096

Kaller, C. P., Unterrainer, J. M., Rahm, B., and Halsband, U. (2004). The impact of problem structure on planning: insights from the Tower of London task. *Cogn. Brain Res.* 20, 462–472. doi: 10.1016/j.cogbrainres.2004.04.002

Katsoni, E., Byrd, D. L., and Casey, B. J. (2006). Special considerations for functional magnetic resonance imaging of pediatric populations. *J. Magn. Reson. Imaging* 23, 877–886. doi: 10.1002/jmri.20578

Kerr, A., and Zelazo, P. D. (2004). Development of "hot" executive function: the children's gambling task. *Brain Cogn.* 55, 148–157. doi: 10.1016/S0278-2626(03)00275-6

Klahr, D., and Robinson, M. (1981). Formal assessment of problem-solving and planning processes in preschool children. *Cogn. Psychol.* 13, 113–148. doi: 10.1016/0010-0285(81)90006-2

Krikorian, R., Bartok, J., and Gay, N. (1994). Tower of London procedure: a standard method and developmental data. *J. Clin. Exp. Neuropsychol.* 16, 840–850. doi: 10.1080/01688639408402697

Lamm, C., Zelazo, P. D., and Lewis, M. D. (2006). Neural correlates of cognitive control in childhood and adolescence: disentangling the contributions of age and executive function. *Neuropsychologia* 44, 2139–2148. doi: 10.1016/j.neuropsychologia.2005.10.013

Lane, R. D., Reiman, E. M., Ahern, G. L., and Thayer, J. F. (2001). Activity in medial prefrontal cortex correlates with vagal component of heart rate variability during emotion: in Rotman Institute Abstracts: Part 1: Abstracts of Presentations. *Brain Cogn.* 49, 185–193.

Lin, F., Heffner, K., Mapstone, M., Chen, D. G. D., and Porsteisson, A. (2013). Frequency of mentally stimulating activities modifies the relationship between cardiovascular reactivity and executive function in old age. *Am. J. Geriatr. Psychiatry* 22, 1210–1221. doi: 10.1016/j.jagp.2013.04.002

Little, R. J. A., and Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.

Luciana, M., and Nelson, C. A. (2002). Assessment of neuropsychological function through the use of the Cambridge neuropsychological testing automated battery: performance in 4- to12-year-old children. *Dev. Neuropsychol.* 22, 595–624. doi: 10.1207/S15326942DN2203_3

Luria, A. R. (1966). *Higher Cortical Functions in Man*. Oxford: Basic Books.

Matthews, S. C., Paulus, M. P., Simmons, A. N., Nelesen, R. A., and Dimsdale, J. E. (2004). Functional subdivisions within anterior cingulate cortex and their relationship to autonomic nervous system function. *Neuroimage* 22, 1151–1156. doi: 10.1016/j.neuroimage.2004.03.005

Mathewson, K. J., Jetha, M. K., Drmic, I. E., Bryson, S. E., Goldberg, J. O., Hall, G. B., et al. (2010). Autonomic predictors of Stroop performance in young and middle-aged adults. *Int. J. Psychophysiol.* 76, 123–129. doi: 10.1016/j.ijpsycho.2010.02.007

McEvoy, L. K., Smith, M. E., and Gevins, A. (1998). Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice. *Cereb. Cortex* 8, 563–574. doi: 10.1093/cercor/8.7.563

Mezzacappa, E., Kindlon, D., Saul, J. P., and Earls, F. (1998). Executive and motivational control of performance task behavior, and autonomic heart-rate regulation in children: physiologic validation of two-factor solution inhibitory control. *J. Child Psychol. Psychiatry* 39, 525–531. doi: 10.1017/S0021963098002273

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Mulder, L. J. M., De Waard, D., and Brookhuis, K. A. (2004). "Estimating mental effort using heart rate and heart rate variability," in *Handbook of Human Factors and Ergonomics Methods*, eds N. Stanton, A. Hedge, H. W. Hendrick, K. A. Brookhuis, and E. Salas (Boca Raton, FL: CRC Press), 666.

Müller, N. G., Mochado, L., and Knight, R. T. (2002). Contributions of subregoins of the prefrontal cortex to working memory: evidence from brain lesions in humans. *J. Cogn. Neurosci.* 14, 673–686. doi: 10.1162/089892902601738582

Nelson, C. A., Monk, C. S., Lin, J., Carver, L. J., Thomas, K. M., and Truwit, C. L. (2000). Functional neuroanatomy of spatial working memory in children. *Dev. Psychol.* 36, 109–116. doi: 10.1037/0012-1649.36.1.109

Newman, S. D., Carpenter, P. A., Varma, S., and Just, M. A. (2003). Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologica* 41, 1668–1682. doi: 10.1016/S0028-3932(03)00091-5

Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59. doi: 10.1002/hbm.20131

Porges, S. W. (2007). The polyvagal perspective. *Biol. Psychol.* 74, 116–143. doi: 10.1016/j.biopsycho.2006.06.009

Porges, S. W., and Byrne, E. A. (1992). Research methods for measurement of heart rate and respiration. *Biol. Psychol.* 34, 93–130. doi: 10.1016/0301-0511(92)90012-J

Porges, S. W., Heilman, K. J., Bazhenova, O. V., Bal, E., Doussard-Roosevelt, J. A., and Koledin, M. (2007). Does motor activity during psychophysiological paradigms confound the quantification and interpretation of heart rate and heart rate variability measures in young children? *Dev. Psychobiol.* 49, 485–494. doi: 10.1002/dev.20228

Royston, P. (1992). Which measures of skewness and kurtosis are best? *Stat. Med.* 11, 333–343. doi: 10.1002/sim.4780110306

Saint-Aubin, J., Klein, R. M., and Landry, T. (2005). Age changes in the missing-letter effect revisited. *J. Exp. Child Psychol.* 91, 158–182. doi: 10.1016/j.jecp.2005.01.007

Shallice, T. (1982). Specific impairments of planning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 298, 199–209. doi: 10.1098/rstb.1982.0082

Speer, N. K., Jacoby, L. L., and Braver, T. S. (2003). Strategy-dependent changes in memory: effects in behavior and brain activity. *Cogn. Affect. Behav. Neurosci.* 3, 155–167. doi: 10.3758/CABN.3.3.155

Stark, C. E. L., and Squire, L. R. (2001). When zero is not zero. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12760–12766. doi: 10.1073/pnas.221462998

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Suess, P. E., Newlin, D. B., and Porges, S. W. (1997). Motivation, sustained attention, and autonomic regulation in school-age boys exposed in utero to opiates and alcohol. *Exp. Clin. Psychopharmacol.* 5, 375–387. doi: 10.1037/1064-1297.5.4.375

Suess, P. E., Porges, S. W., and Plude, D. J. (1994). Cardiac vagal tone and sustained attention in school-age children. *Psychophysiology* 31, 17–22. doi: 10.1111/j.1469-8986.1994.tb01020.x

Thatcher, R. W., North, D. M., and Biver, C. J. (2008). Development of cortical connections as measured by EEG coherence and phase delays. *Hum. Brain Mapp.* 29, 400–415. doi: 10.1002/hbm.20474

Thayer, J. F., Hansen, A. L., Saus-Rose, E., and Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. *Ann. Behav. Med.* 37, 141–153. doi: 10.1007/s12160-009-9101-z

Thayer, J. F., and Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *J. Affect. Disord.* 61, 201–216. doi: 10.1016/S0165-0327(00)00338-4

Thomas, K. M., King, S. W., Frazen, P. L., Welsh, T. F., Berkowitz, A. L., Noll, D. C., et al. (1999). A developmental fMRI study of spatial working memory. *Neuroimage* 10, 327–338. doi: 10.1006/nimg.1999.0466

Unterrainer, J. M., Kaller, C. P., Loosli, S. V., Heinze, K., Ruh, N., Paschke-Müller, M., et al. (2014). Looking ahead from age 6 to 13: a deeper insight into the development of planning ability. *Br. J. Psychol.* 106, 46–67. doi: 10.1111/bjop.12065

Unterrainer, J. M., Rahm, B., Kaller, C. P., Leonhart, R., Quiske, K., Hoppe-Seyler, K., et al. (2004a). Planning abilities and the Tower of London: is this task measuring a discrete cognitive function? *J. Clin. Exp. Neuropsychol.* 26, 846–856. doi: 10.1080/13803390490509574

Unterrainer, J. M., Rahm, B., Kaller, C. P., Ruff, C. C., Spreer, J., Krause, B. J., et al. (2004b). When planning fails: individual differences and error-related brain activity in problem solving. *Cereb. Cortex* 14, 1390–1397. doi: 10.1093/cercor/bhh100

Vuontela, V., Steenari, M. R., Aronen, E. T., Korvenoja, A., Aronen, H. J., and Carlson, S. (2009). Brain activation and deactivation during location and color working memory tasks in 11-13-year old children. *Brain Cogn.* 69, 56–64. doi: 10.1016/j.bandc.2008.05.010

Vuontela, V., Steenari, M. R., Carlson, S., Koivisto, J., Fjällberg, M., and Aronen, E. (2003). Audiospatial and visuospatial working memory in 6- to 13-year-old school children. *Learn. Mem.* 10, 74–81. doi: 10.1101/lm.53503

Welsh, M. C., Pennington, B. F., and Groisser, D. B. (1991). A normative-developmental study of executive function: a window on prefrontal function in children. *Dev. Neuropsychol.* 7, 131–149. doi: 10.1080/87565649109540483

Zelazo, P. D. (2000). "Self-reflection and the development of consciously controlled processing," in *Children's Reasoning and the Mind,* eds P. Mitchell and K. J. Riggs (Hove: Psychology Press/Taylor & Francis), 169–189.

Zelazo, P. D., and Müller, U. (2007). "Executive function in typical and atypical development," in *Blackwell Handbook of Childhood Cognitive Development,* ed. U. Goswami (Malden, MA: Blackwell Publishers Ltd).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 January 2014; accepted: 30 November 2014; published online: 05 March 2015.

Citation: Byrd DL, Reuther ET, McNamara JPH, DeLucca TL and Berg WK (2015) Age differences in high frequency phasic heart rate variability and performance response to increased executive function load in three executive function tasks. Front. Psychol. ***5****:1470. doi: 10.3389/fpsyg.2014.01470*

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2015 Byrd, Reuther, McNamara, DeLucca and Berg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 13 March 2014
doi: 10.3389/fpsyg.2014.00218

Task-evoked pupillometry provides a window into the development of short-term memory capacity

Elizabeth L. Johnson[1,2], Alison T. Miller Singley[1,2], Andrew D. Peckham[1], Sheri L. Johnson[1] and Silvia A. Bunge[1,2]*

[1] Department of Psychology, University of California Berkeley, Berkeley, CA, USA
[2] Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA, USA

Edited by:
Nicolas Chevalier, University of Edinburgh, UK

Reviewed by:
Jenny Richmond, University of New South Wales, Australia
Kimberly Sarah Chiew, Duke University, USA

****Correspondence:***
Silvia A. Bunge, Department of Psychology, University of California, Berkeley, 3407 Tolman Hall, Berkeley, CA 94720-1650, USA
e-mail: sbunge@berkeley.edu

The capacity to keep multiple items in short-term memory (STM) improves over childhood and provides the foundation for the development of multiple cognitive abilities. The goal of this study was to measure the extent to which age differences in STM capacity are related to differences in task engagement during encoding. Children ($n = 69$, mean age = 10.6 years) and adults ($n = 54$, mean age = 27.5 years) performed two STM tasks: the forward digit span test from the Wechsler Intelligence Scale for Children (WISC) and a novel eyetracking digit span task designed to overload STM capacity. Building on prior research showing that task-evoked pupil dilation can be used as a real-time index of task engagement, we measured changes in pupil dilation while participants encoded long sequences of digits for subsequent recall. As expected, adults outperformed children on both STM tasks. We found similar patterns of pupil dilation while children and adults listened to the first six digits on our STM overload task, after which the adults' pupils continued to dilate and the children's began to constrict, suggesting that the children had reached their cognitive limits and that they had begun to disengage from the task. Indeed, the point at which pupil dilation peaked at encoding was a significant predictor of WISC forward span, and this relationship held even after partialing out recall performance on the STM overload task. These findings indicate that sustained task engagement at encoding is an important component of the development of STM.

Keywords: short-term memory, digit span, task-evoked pupillary response, pupillometry, development

INTRODUCTION

The ability to maintain information for a short period of time, known variably as short-term memory (STM) or the storage component of working memory, increases over childhood (for meta-analysis see Simmering and Perone, 2013). STM capacity is tied to the ability to perform complex cognitive tasks, such as reading and math (Baddeley, 1992; Cowan et al., 2011), and the development of STM capacity partially governs age-related gains in higher-order cognitive functions (Bayliss et al., 2005; Magimairaj and Montgomery, 2012). The goal of the present study was to gain mechanistic insights into developmental changes and individual differences in STM capacity.

One of the most commonly used indices of STM in children is the digit span task, a measure of verbal STM (Bayliss et al., 2005; Cowan et al., 2005). The digit span task requires the encoding and immediate serial recall of a list of numbers presented aurally, and the length of an individual's span depends on how well s/he can attend to, rehearse, and subsequently repeat back the stimuli. The ability to remember long lists in simple span tasks has been validated as a robust correlate of higher-order cognitive functions as measured by complex span tasks in children (Cowan et al., 2005) and adults (Unsworth and Engle, 2007a,b). Age-related changes and individual differences in digit span could in theory reflect differences in cognitive resource allocation at encoding, rehearsal, and/or recall. Here, we sought to assess the extent to which age-related changes and individual differences in STM capacity could be explained by differences in cognitive effort during stimulus encoding, as measured via the task-evoked pupillary response to cognitive load (Hess and Polt, 1964; Beatty, 1982; Beatty and Lucero-Wagoner, 2000; Karatekin, 2007; Laeng et al., 2012).

Pupil size is governed both by ambient light levels and physiological arousal (Kahneman, 1973; Beatty, 1982; Beatty and Lucero-Wagoner, 2000; Karatekin, 2007; Laeng et al., 2012). Pupil dilation related to physiological arousal is mediated by the simultaneous activation of sympathetic pathways and inhibition of parasympathetic pathways (Beatty and Lucero-Wagoner, 2000), and evidence suggests that task-evoked pupil dilation results from cortical inhibition of the parasympathetic oculomotor nucleus (Wilhelm et al., 1999; Steinhauer et al., 2004). During a state of heightened attention, neurons in the locus coeruleus fire rapidly, supplying high levels of noradrenaline to numerous targets throughout the body, including both the eyes and brain. In the eye, this neurotransmitter mediates pupil dilation; in the brain, it regulates attention through its modulatory effects on brain activity (see Gilzenrat et al., 2010; Laeng et al., 2012; Donner and Nieuwenhuis, 2013; Eldar et al., 2013).

Task-evoked pupil dilation in well-controlled experimental settings has been referred to variably as a peripheral marker of heightened attention, mental effort, or allocation of cognitive

control when the task prompts focus or conscious engagement. Kahneman (1973) described it as reflecting the "intensive aspect" of attention; more recently, Gilzenrat et al. (2010) have described task-evoked pupillary dilation as reflecting task engagement. Indeed, a large body of research provides compelling evidence that task-evoked pupil dilation is sensitive to cognitive load (Beatty, 1982; Beatty and Lucero-Wagoner, 2000). Beginning with Kahneman and Beatty (1966), researchers have consistently shown that adults' pupils dilate incrementally with each digit encoded in a digit span task until the length of the digit sequence exceeds STM capacity, at which point pupil size begins to plateau or diminish (Kahneman et al., 1968; Peavler, 1974; Granholm et al., 1996, 1997; Cabestrero et al., 2009). Pupils also tend to constrict during recall as items are offloaded from STM (Kahneman and Beatty, 1966; Cabestrero et al., 2009). These findings are consistent with the idea that cognitive resources are dedicated in a manner proportionate to the cognitive load.

Pupil dilation patterns have also been used to examine individual differences in cognitive functioning among adults. Ahern and Beatty (1979, 1981) showed that cognitively higher-functioning adults—as defined based on their scores on the Scholastic Aptitude Test—exhibited consistently smaller dilation amplitudes on STM, mental multiplication, and sentence comprehension tasks than lower-functioning adults. These patterns of pupil dilation were interpreted as indices of mental effort, suggesting that performance of the same cognitive task was less challenging for higher-functioning adults. Taken together, the results of prior studies validate pupil dilation as a measure of task engagement, with pupils dilating as cognitive effort is expended.

Simmering and Perone (2013) have argued that the field of cognitive development would benefit from research linking theory to real-time behavior; specifically, they call for approaches that combine evidence from "micro-behavior"—i.e., indices of mechanisms underlying cognitive processes—and "macro" measures such as performance accuracy. We propose that task-evoked pupillometry represents a "micro" index of mental effort that can be used to probe developmental changes in task engagement. Given its high temporal resolution, well-validated use in studies of adult cognition, and non-invasive nature, task-evoked pupillometry has the potential to provide important insights with regard to cognitive development (cf. Karatekin, 2007; Laeng et al., 2012).

Thus far, there have been only a few studies of task-evoked pupillometry involving children (Boersma et al., 1970; Karatekin, 2004, 2007; Karatekin et al., 2007a,b; Chatham et al., 2009), and only one of these studies involved a digit span task (Karatekin, 2004). In this study, 10-year-olds ($n = 15$) and young adults ($n = 21$) performed a digit span task in which they listened to sequences of 4, 6, and 8 digits. Although the 10-year-olds did not perform as well as the adults on either the 6- or 8-digit sequences, their patterns of pupil dilation differed only when they encoded the 8-digit sequences (Karatekin, 2004). On these long sequences, children exhibited shallower mean rates of dilation per digit than did adults, which the authors interpreted as indicating that they allocated fewer cognitive resources to the task.

Here, we sought to more closely examine the relationships between task engagement at encoding and developmental changes and individual differences in STM capacity. To this end, we measured pupil diameter continuously as participants encoded digit sequences that exceeded typical STM capacity, i.e., an STM overload task. If, as the results of Karatekin (2004) suggest, children are unable to recruit cognitive resources sufficient to encode at high loads, then their pupils should stop dilating (Cabestrero et al., 2009) and/or constrict (Peavler, 1974; Granholm et al., 1996) earlier in the sequence as compared to adults. Seeking to explore the relationship between these task-evoked pupillary responses and differences in STM capacity, we also administered the forward span task from the Digit Span subtest of the Wechsler Intelligence Scale for Children (Wechsler, 2003) to both children and adults. We hypothesized that if the point at which pupil diameter asymptotes is related to the amount of information encoded into STM, then this value should be related to STM capacity.

METHODS

PARTICIPANTS

Sixty-nine healthy children (36 males, 33 females; ages 7.5–14.0 years, mean 10.6 ± 1.1 years) and 54 healthy adults (27 males, 27 females; ages 18.3–60.8 years, mean 27.5 ± 10.8 years) participated in this study.[1] Children were recruited through the Berkeley Chess School outreach program at public schools in Oakland, CA, or surrounding San Francisco Bay Area communities, and thanked via a classroom gift by request of the school administration. Adults were recruited from the University of California, Berkeley, or the San Francisco Bay Area via advertisements, and received monetary compensation or—for UC Berkeley students in the Research Participation Pool—course credit. All participants had normal or corrected-to-normal vision and hearing, and were fluent in English.

BEHAVIORAL FORWARD DIGIT SPAN

To assess STM capacity, we used the forward span task in the Digit Span subtest on the Wechsler Intelligence Scale for Children—Fourth Edition (WISC-IV; Wechsler, 2003). The forward span task is a commonly used behavioral measure of verbal STM in multiple populations (Kane et al., 2004; Bayliss et al., 2005; Cowan et al., 2005; Alloway et al., 2009). The Digit Span subtest procedure is identical in the children and adult Wechsler test batteries; we chose to use the WISC subtest across age groups to keep the digit lists constant. Participants are read a series of digits (e.g., "9, 4, 2") at a rate of one digit per second and are asked to repeat the digits back to the experimenter in the same serial order presented. Two trials are presented at each span length, starting with two digits per trial. If the participant repeats at least one of the two trials of the same sequence length successfully, the experimenter presents two trials of a sequence that is one digit longer. This procedure continues until the participant misses both trials of a particular span length or completes the trials with the maximum 9-digit span.

[1] Three adults and one child who reported having taken medications on the day of testing were excluded from the current sample. Two adults took an antihistamine and one took Flomax; the child's medication is not known. Six of the young adults recruited through the UC Berkeley Research Participant Pool did not provide their exact ages.

In tests of verbal STM, healthy adults remember an average of seven digits, plus or minus two (Miller, 1956); children tend to remember fewer digits than adults (Simmering and Perone, 2013). An individual's STM span is calculated as the length of the longest sequence of digits successfully repeated back to the experimenter, for a maximum of 9. The forward total score reflects the number of trials each participant completed correctly, for a maximum of 16.

STM OVERLOAD TASK

Following administration of the WISC forward digit span, participants completed a computerized STM overload task while undergoing eyetracking. Our task was adapted from Peavler (1974), Granholm et al. (1996, 1997), Karatekin (2004), and Cabestrero et al. (2009). As in the WISC task, participants heard a sequence of digits, presented at the rate of one digit per second, and were asked to repeat them back immediately in the same order presented (Wechsler, 2003). In our adaptation of the task, participants completed a total of four trials, all involving the same number of digits. Children were asked to encode sequences of nine digits, whereas adults were asked to encode sequences of 11 digits (the same nine digits as for the children, with two additional digits added at the end of the sequence). These digit sequence lengths were chosen because they exceed average WISC forward spans, allowing us to examine pupillary responses once participants surpassed their individual encoding limitations (Granholm et al., 1996, 1997; Karatekin, 2004; Cabestrero et al., 2009). For the present purposes, we were interested in average pupil dilation and subsequent serial recall accuracy for each digit.

All participants were informed that they would hear a series of numbers. They were instructed to remember the digits as presented and then do their best to recall the full sequence of digits in the correct order. Each trial began with a 1-s auditory cue ("memorize"), alerting participants to the beginning of a trial. After the last digit for the trial was presented, the word "recall" signaled the participant to repeat the numbers back; as in the WISC forward digit span, the recall phase was self-paced. Participants completed all four trials irrespective of recall accuracy. The experimenter manually recorded participants' responses during the recall phase.

Both children and adults completed the same two practice trials before the experimental trials: a 3-digit trial followed by a 5-digit trial. They were permitted to repeat this round by request. After practice, participants underwent a 5-point eyetracking calibration procedure, and then began the experimental trials. Within each age group, all participants completed the same four experimental trials, with the order of trials randomized.

Participants were instructed to look at a 1×1 inch fixation cross in the middle of the screen, presented in white on a black background, throughout the computer task. This design permitted the recording of pupil data at fixed luminance for the duration of the task, ensuring that pupillary responses were independent of pupillary light reflexes (Beatty, 1982; Beatty and Lucero-Wagoner, 2000). To allow participants' pupil diameters to return to a neutral baseline before the start of each trial (e.g., Cabestrero et al., 2009; van der Meer et al., 2010), we programmed the task in such a way that it proceeded automatically to the next trial only after the eyetracker had captured 2 s of continuous data.

EYETRACKING APPARATUS

Stimuli were presented using the Tobii E-Prime Software Extensions (Psychology Software Tools, Pittsburgh, PA), which syncs the timing of stimulus presentation with a second computer that records pupil data. Participants were seated comfortably in front of the Tobii T120 Eye Tracker (17-inch monitor, 1280×1024 pixel resolution); distance was calibrated individually so that each participant focused on the middle of the screen, within a range of 50–80 cm. The Tobii T120 built-in camera captures data with a temporal resolution of 120 Hz, producing a data point every 8.3 ms, and average spatial resolution of 0.3° of visual angle. Because the camera can automatically compensate for small head movements (within a 30×22 cm area at 70 cm distance), participants' heads were not restrained. The camera simultaneously recorded the pupil diameter of the left and right eyes.

DATA ANALYSES

Nineteen children and eight adults were excluded from the sample due to insufficient recording of eyetracking data, yielding data from 69 children and 54 adults. We considered recordings insufficient if pupil data were absent across all four trials of at least one digit or while hearing the "memorize" cue (i.e., the cue period), or if less than 25% of data remained overall after cleaning the data to remove artifacts (adapted from Granholm et al., 1996; Siegle et al., 2011). These were cases of either technical error or excessive blinking or head motion on the participant's part, and so using such stringent cutoffs permitted us to perform analyses without need for interpolating data points to fill gaps in data collection.

Data were cleaned using a local fit procedure. We manually inspected graphic displays of a subset of data in each group sample for artifacts (e.g., partial eyelid closures, apparent changes in diameter resulting from motion), and then implemented a computer algorithm to automate this process for all subjects. A local regression model was applied to the full datasets (loess model; Cleveland et al., 1992), such that data points were removed from analysis if they fell out of the range of five standard errors above or below the locally defined, weighted mean. We applied this process separately to the raw pupil diameter of each eye, fitting locally over 400-ms segments of data around each diameter data point.[2] Because subjects' heads were not restrained, we also applied this procedure to the mean distance between subjects' eyes and the camera. We used a more conservative fit based on 200 ms around each distance data point in order to pick up artifacts due to abrupt changes in head position. Overall, data were discarded if they fell out of range in either eye based on pupil diameter, or based on distance; fewer than 4% of data points were removed in this procedure.

To measure pupil dilation during encoding, we calculated the average pupil diameter across both eyes at each remaining data point (8.3 ms). Data for one eye were used when data for both were not available. We then calculated the mean diameter over each second, time-locked to the presentation of each

[2] A wider range of data points, up to 700 ms on pupil diameter and 500 ms on distance, was used on datasets with fewer recorded data points, as required by the loess model.

stimulus, averaged across the four experimental trials. This procedure yielded one data point for the "memorize" cue, and either nine or eleven data points for the digit sequence, depending on whether the participant was a child or an adult.

The absolute diameter of the pupil at rest is known to decrease from childhood into adulthood. This age-related change is posited to reflect a gradual decrease over childhood in the influence of the sympathetic branch concurrent with a decrease in central inhibition of the parasympathetic pathway (Karatekin et al., 2007a). Thus, to compare patterns of pupil dilation between children and adults, it is necessary to control for these differences in baseline pupil diameter.

Task-evoked pupil dilation was defined as the percentage of dilation at each digit, over 1 s, relative to the mean pupil diameter over the 1-s cue period, i.e., $\text{dilation}_{\text{digit}} = (\text{diameter}_{\text{digit}} - \text{diameter}_{\text{cue}}) / \text{diameter}_{\text{cue}}$ (Karatekin, 2004; also Hess and Polt, 1964; Beatty and Lucero-Wagoner, 2000). Pupil dilation data were submitted to a mixed-model, repeated-measures analysis of variance (ANOVA), with digit as the within-subjects factor and age group as the between-subjects factor. Planned *post-hoc* comparisons between dilation at each digit and the next consecutive digit in the sequence were performed within each age group.

Recall accuracy was defined as the proportion of digits correctly recalled as a function of serial position on the STM overload task (Cowan et al., 2005). If a participant correctly recalled the first digit on all four trials, s/he was given an accuracy of 1 on the first digit. If, however, a participant correctly recalled a digit on three of the four trials, and missed it or recalled it incorrectly on one trial, s/he was given an accuracy of 0.75 for that digit. This procedure yielded values of 1, 0.75, 0.5, 0.25, or 0 for each digit. We conducted a mixed-model, repeated-measures ANOVA, and performed *post-hoc* comparisons between each digit and the next digit in the sequence within each age group. We also conducted regression analyses to further explore the relationships between measures of STM capacity and pupillary dilation at encoding, controlling for age group.

RESULTS

AGE-RELATED DIFFERENCES IN STM

First, we tested for group differences in STM capacity on the WISC digit span test and on our computerized STM overload test. As expected, adults had significantly higher WISC forward spans and scores than children, $t_{\text{span}(115.1)} = 7.6$, $t_{\text{score}(117.8)} = 7.9$; both $p < 0.001$ (**Table 1**). On our STM overload task, adults recalled more digits than children (**Figure 1**, **Table 1**). A 9 (digit: 1 through 9) × 2 (age group) ANOVA revealed significant main effects of digit, $F_{(8,\ 960)} = 258.92$, $MSE = 0.03$, $p < 0.001$, $\eta^2 = 0.68$, and age group, $F_{(1,\ 120)} = 68.87$, $MSE = 0.15$, $p < 0.001$, $\eta^2 = 0.37$.

Both groups exhibited a primacy effect, such that proportion of correctly recalled digits was high at the beginning of the digit sequence and diminished with each additional digit (i.e., serial position), consistent with prior research on immediate serial recall (Kane et al., 2004; Unsworth and Engle, 2007a,b). In adults, there were significant incremental decreases from positions 1 to 2, 2 to 3, 3 to 4, 6 to 7, 7 to 8, 8 to 9, and 9 to 10 [all $t_{(53)} > 3.0$, $p < 0.01$]; and in children, from positions 1 to 2, 3 to 4, 6 to 7, 7

Table 1 | Descriptive statistics for WISC, pupillary, and recall accuracy data by age group.

	Adults	Children	Group differences
	Mean (*SD*)	**Mean (*SD*)**	
WISC FORWARD SPAN TASK			
Span	7.19 (1.23)	5.47 (1.26)	$t_{(115.11)} = 7.56$, $p < 0.001$
Total score	11.41 (2.12)	8.24 (2.33)	$t_{(117.84)} = 7.86$, $p < 0.001$
STM OVERLOAD TASK			
Encoding phase			
Mean pupil diameter in mm			
Cue	3.81 (0.57)	3.99 (0.47)	
Digit 1	3.88 (0.60)	4.06 (0.48)	
Digit 2	3.93 (0.63)	4.08 (0.52)	
Digit 3	3.94 (0.62)	4.13 (0.51)	
Digit 4	3.98 (0.61)	4.15 (0.53)	
Digit 5	4.01 (0.62)	4.19 (0.51)	
Digit 6	4.09 (0.63)	4.21 (0.51)	
Digit 7	4.12 (0.64)	4.20 (0.52)	
Digit 8	4.14 (0.65)	4.17 (0.51)	
Digit 9	4.14 (0.66)	4.15 (0.54)	
Digit 10	4.15 (0.66)	n/a	
Digit 11	4.13 (0.67)	n/a	
Digit-at-peak dilation	7.65 (1.81)	6.10 (2.02)	$t_{(118.73)} = 4.46$, $p < 0.001$
Recall phase			
Total correct	4.79 (1.35)	2.90 (1.13)	
Proportion correct	0.44 (0.12)	0.32 (0.13)	$t_{(114.82)} = 4.99$, $p < 0.001$

WISC and recall phase data were missing for one child. Digit-at-peak dilation computations are based on data from digits 1 to 9. Independent-samples t-tests were performed on variables that were standardized for comparison across age groups.

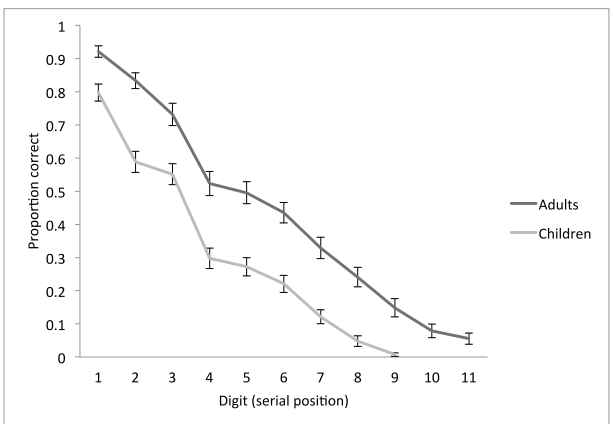


FIGURE 1 | Behavioral performance on the STM overload task. Mean proportion of digits correctly recalled as a function of serial position, plotted separately for children and adults. Error bars represent standard mean error.

to 8, and 8 to 9 [all $t_{(67)} > 2.7$, $p < 0.01$]. A follow-up one-way ANOVA showed that adults were significantly more accurate than children on all digits, all $p < 0.001$, and an independent samples *t*-test confirmed that adults recalled 12% more digits than children overall ($p < 0.001$, see **Table 1**). This finding is consistent with prior literature on the development of STM, showing that capacity increases with age from childhood into adulthood (Simmering and Perone, 2013).

Next, we used partial correlation analyses to test whether the standardized WISC digit span subtest and our STM overload task elicited similar behaviors, controlling for age group. This analysis showed that recall accuracy on the STM overload task was significantly, albeit modestly, correlated with WISC score after controlling for group [$r_{(119)} = 0.19, p < 0.04$]. The partial correlation between recall accuracy and WISC span, however, did not retain significance [$r_{(119)} = 0.14, p < 0.12$].

These findings suggest that the cognitive factors that contribute to performance on our STM overload task overlap partially with those of the standard digit span task, in which the length of the test sequence increases only after mastery is demonstrated at a particular sequence length. Indeed, behavioral performance on a memory test reflects the combined outcome of cognitive processes operating during encoding, maintenance, and retrieval. Given the high temporal resolution of pupillometry, by contrast, it is possible to examine measurements taken during a specific task phase. Here, we probe the relationships between STM capacity and pupil dilation during the encoding phase of our STM overload task.

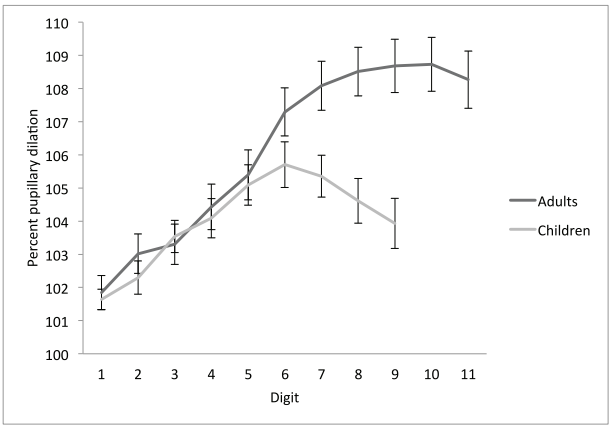


FIGURE 2 | Temporal dynamics of pupil dilation and constriction on the STM overload task. Mean percentage of pupil dilation for each digit relative to mean pupil diameter over the cue period (set to a starting point of 100%; Karatekin, 2004), by age group. Adults encoded four sequences of 11 digits each, and children encoded four sequences of 9 digits each. Error bars represent standard mean error.

AGE-RELATED DIFFERENCES IN PUPIL DILATION AT ENCODING

In accordance with our research aim of investigating the relationship between task-evoked pupillary responses and STM capacity, we tested for group differences in dilation relative to the cue period immediately prior to task. Consistent with prior work (Karatekin, 2004; also Beatty and Lucero-Wagoner, 2000), children had larger pupils at all timepoints than adults (**Table 1**); thus, we plotted pupil dilation in terms of percentage change from the cue period (**Figure 2**).

A 9 (digit) × 2 (age group) ANOVA revealed significant main effects of digit, $F_{(8, 968)} = 59.24$, $MSE = 7.23$, $p < 0.001$, $\eta^2 = 0.33$, and age group, $F_{(1, 121)} = 4.09$, $MSE = 168.03$, $p < 0.05$, $\eta^2 = 0.03$, and a significant digit × group interaction, $F_{(8, 968)} = 13.51$, $MSE = 7.23$, $p < 0.001$, $\eta^2 = 0.10$. Both age groups demonstrated an increase in pupil dilation as a function of digit, to a point. Adults' pupils showed incremental increases from cue to digit 1, and digits 1 to 2, 3 to 4, 4 to 5, 5 to 6, and 6 to 7 [all $t_{(53)} > 2.8$, $p < 0.01$], and continued to dilate until almost digit 9 on average (8.7 ± 2.2). Children's pupils dilated until digit 6 on average (6.1 ± 2.0), with incremental increases from cue to digit 1, digit 2 to 3, and digit 4 to 5, [all $t_{(68)} = 2.7$, $p < 0.01$], and a marginally significant increase from digit 1 to 2 [$t_{(68)} = 2.0$, $p = 0.05$]. In contrast, a significant decrease was observed from digit 7 to 8, $t_{(68)} = 2.1$, $p < 0.05$.

A one-way ANOVA with age group as the between-subjects factor confirmed that adults' pupils were significantly more dilated than children's while encoding digits 7, 8, and 9 (all $p < 0.01$), indicating that where adults' pupil diameters continued to dilate or reached a stable plateau, children's pupils reached an asymptote or began to constrict. The age groups did not differ significantly in pupil dilation on digits 1 through 6 (all $p > 0.12$), suggesting a similar rate of dilation within the constraints of STM capacity.

To directly compare the latency to peak pupil dilation—i.e., digit-at-peak—between groups, we also conducted a planned comparison based on the digit (1–9) at which pupils reached maximum dilation. Adults' maximum pupil dilation occurred on average at digit 7.7 ± 1.8, which was significantly greater than children's maximum at digit 6.1 ± 2.0, $t_{(118.7)} = 4.5$, $p < 0.001$ (**Table 1**).

RELATIONSHIPS BETWEEN PUPIL DILATION AND STM

The correspondence between average digit-at-peak values (7.7 and 6.1 for adults and children, respectively) and average WISC spans (7.2 and 5.5) hints at a relationship between STM capacity and the dynamics of pupil dilation during STM encoding. To test this hypothesis directly, we first conducted linear regression analyses between the pupillary measure of digit-at-peak dilation and each behavioral STM measure: recall accuracy on the STM overload task, WISC span, and WISC score. Digit-at-peak was significantly correlated with all three measures ($\beta_{recall} = 0.30$, $\beta_{span} = 0.38$, $\beta_{score} = 0.37$; all $p \leq 0.001$). The correlation between digit-at-peak and each WISC measure retained significance after partialing out recall accuracy on the STM overload task [$r_{span(119)} = 0.30$, $r_{score(119)} = 0.29$; both $p = 0.001$].

Next, we measured the extent to which individual variability in digit-at-peak explained individual differences in STM capacity, controlling for age group. In a multiple regression analysis, we modeled STM capacity as a function of digit-at-peak and group. This analysis revealed a strong effect of group on all three STM measures, as expected, as well as an independent contribution of digit-at-peak to each measure, $p < 0.05$ (see **Table 2** for full

Table 2 | Multiple regression analyses for WISC score, WISC span, and recall accuracy

	B	*SE B*	β
WISC FORWARD SPAN			
Digit-at-peak	0.14	0.06	0.19*
Group	−1.50	0.24	−0.50**
WISC FORWARD TOTAL SCORE			
Digit-at-peak	0.24	0.10	0.18*
Group	−0.28	0.43	−0.51**
STM OVERLOAD RECALL ACCURACY			
Digit-at-peak	0.01	0.01	0.18*
Group	−0.10	0.02	−0.35**

*$p < 0.05$, **$p < 0.001$

results). These results indicate that cognitive resource allocation at encoding, as measured by the point of maximal pupil dilation on our STM overload task, can explain individual differences in STM capacity on a standard digit span task.

DISCUSSION

Consistent with decades of prior research in adults, the present results corroborate a close link between cognitive demands imposed by the digit span task and task-evoked pupil dilation (Kahneman and Beatty, 1966; Kahneman et al., 1968; Peavler, 1974; Granholm et al., 1996, 1997; Cabestrero et al., 2009), and show that children also exhibit this link (also Karatekin, 2004). Our findings extend prior work in two ways. First, we provide evidence that the children disengaged from the task as soon as the cognitive load surpassed their STM capacity, whereas adults stayed engaged while encoding additional items beyond their span. Second, we show that the point at which pupil dilation peaks is related to STM capacity—independent of age, and even after partialing out recall accuracy on the STM overload task.

With our STM span overload paradigm, we obtained similar trajectories of pupil dilation for children and adults until the sixth digit, after which the age groups diverged. Whereas adults showed dilation during encoding up to the ninth digit and then exhibited a plateau in pupil diameter until the end of the 11-digit sequence, children's pupils plateaued from digit 6 to 7, constricted from 7 to 8, and then plateaued until the end of the 9-digit sequence. In contrast to Karatekin (2004), who showed that children exhibited shallower dilation than adults during encoding of an 8-digit sequence, this finding shows children and adults dilate at similar rates up to digit 6, after which the groups' dilation patterns diverge.

Analyses focused on digit-at-peak revealed a significant relationship between the ordinal number corresponding to the digit at which maximal pupil dilation was reached on digits 1–9 and STM capacity, as reflected in our STM task and the WISC Digit Span subtest. That is, individual children or adults whose pupils peaked later in the encoded sequence were more likely to have a higher STM span, as reflected in multiple measures. This pupil-behavior relationship, observed independently of age group, is all the more noteworthy because performance on our STM overload task was not significantly related to WISC forward span after partialing out the effect of group. Thus, pupillometry reveals a relationship between encoding on one task and recall on another that would not have been detected via comparison of behavioral performance on the two tasks. These findings suggest that the allocation of cognitive resources—what Kahneman (1973) called the "intensive aspect" of attention—during encoding of information at high cognitive loads is an important contributor to the development of STM.

However, the group difference in STM performance suggests that attention is not the only factor. The groups exhibited the same rate of dilation for digits 1 through 6, indicating a similar level of cognitive effort on those digits, yet adults outperformed children on recall for all digits, not just digits 7 and higher. Thus, similar levels of cognitive resource allocation in children and adults could not fully account for the group difference in recall performance (also Karatekin, 2004). Success on the digit span task requires participants to maintain encoded digits in STM while additional digits are presented, as well as during the recall phase. Attention, echoic memory, rehearsal, and mnemonic strategies are all components of maintenance that contribute to STM performance, and it is likely that each of these cognitive components contributes to the more global measure offered by the task-evoked pupillary response. Further, STM capacity is operationalized in the digit span task as the number of digits that one can accurately recall in the right order via verbal report. This number is likely to be smaller than the number of digits in a sequence that one could accurately identify as "old" on a test of recognition memory (e.g., Unsworth and Engle, 2007b). Pupillometry has been employed in the context of long-term recognition memory (for review see Goldinger and Papesh, 2012), and given the relationship we have found between peak pupil dilation and STM span, it would be of interest to examine how the dynamics of pupil dilation and constriction at encoding relate to subsequent recognition memory as well as recall.

In summary, this study provides insight into the unique relationship between task engagement at encoding and STM capacity, and highlights the role that pupillometry can play in elucidating developmental changes and individual differences in cognition. This work supports Simmering and Perone's (2013) thesis that measures of "micro-behaviors" combined with "macro" performance measures can inform research on cognitive development. Our results further highlight the potential of pupillometry to address inquiries that extend well beyond the study of prototypical adult cognition.

The methodological approach reported here also has practical applications. Our STM overload task could provide insights regarding the cognitive deficits observed in specific patient populations (e.g., in amnesics, Laeng et al., 2007)—or, perhaps in the future, in individual patients. More generally, the task-evoked pupillary response could in theory be used to evaluate the effectiveness of a targeted cognitive intervention, pinpointing precisely at what stage(s) of a task the intervention influences cognitive processing.

AUTHOR NOTE

This study was supported by a James S. McDonnell Foundation Scholar Award and a Tourette Syndrome Association research

grant to Silvia A. Bunge. The Bunge lab's first eyetracking study could not have been completed without the help of numerous individuals. We thank Jordan Tharp and Galen Mancino for assistance with programming and set-up of the eyetracker, Jesse Niebaum, Jordan Tharp, and Farida Valji for assistance collecting and compiling data, Se Ri (Sally) Bae for writing up preliminary results for an undergraduate honor's thesis, Dr. Robert DiMartino for helpful discussions, and Marcus Stoiber and Dr. Davide Risso for invaluable consultation regarding data preprocessing.

REFERENCES

Ahern, S. K., and Beatty, J. (1979). Physiological signs of information processing vary with intelligence. *Science* 205, 1289–1292. doi: 10.1126/science.472746

Ahern, S. K., and Beatty, J. (1981). "Physiological evidence that demand for processing capacity varies with intelligence," in *Intelligence and Learning*, eds M. Friedman, J. P. Dos, and N. O'Connor (New York, NY: Plenum Press), 121–128.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., and Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Dev.* 80, 606–621. doi: 10.1111/j.1467-8624.2009.01282.x

Baddeley, A. D. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359

Bayliss, D., Jarrold, C., Baddeley, A. D., Gunn, D., and Leigh, E. (2005). Mapping the developmental constraints on working memory span performance. *Dev. Psychol.* 41, 579–597. doi: 10.1037/0012-1649.41.4.579

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* 91, 276–292. doi: 10.1037/0033-2909.91.2.276

Beatty, J., and Lucero-Wagoner, B. (2000). "The pupillary system," in *Handbook of Psychophysiology, 2nd Edn.*, eds J. Cacioppo, L. Tassinary, and G. Berntson (Cambridge, UK: Cambridge University Press), 142–162.

Boersma, F., Wilton, K., Barham, R., and Muir, W. (1970). Effects of arithmetic problem difficulty on pupillary dilation in normals and educable retardates. *J. Exp. Child Psychol.* 9, 142–155. doi: 10.1016/0022-0965(70)90079-2

Cabestrero, R., Crespo, A., and Quiros, P. (2009). Pupillary dilation as an index of task demands. *Percept. Mot. Skills* 109, 664–678. doi: 10.2466/pms.109.3.664-678

Chatham, C. H., Frank, M. J., and Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5529–5533. doi: 10.1073/pnas.0810002106

Cleveland, W. S., Grosse, E., and Shyu, W. M. (1992). "Local regression models," in *Statistical Models in S*, eds J. M. Chambers and T. J. Hastie (Pacific Grove, CA: Wadsworth and Brooks/Cole), 309–376.

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cogn. Psychol.* 51, 42–100. doi: 10.1016/j.cogpsych.2004.12.001

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., Gilchrist, A. L., and Scott Saults, J. S. (2011). "New insights into an old problem: distinguishing storage from processing in the development of working memory," in *Cognitive Development and Working Memory: a Dialogue Between neo-Piagetian and Cognitive Approaches*, eds V. Gaillard and P. Barrouillet (Hove; New York: Psychology Press), 137–150.

Donner, T. H., and Nieuwenhuis, S. (2013). Brain-wide gain modulation: the rich get richer. *Nat. Neurosci.* 16, 989–990. doi: 10.1038/nn.3471

Eldar, E., Cohen, J. D., and Niv, Y. (2013). The effects of neural gain on attention and learning. *Nat. Neurosci.* 16, 1146–1153. doi: 10.1038/nn.3428

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., and Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cogn. Affect. Behav. Neurosci.* 10, 252–269. doi: 10.3758/CABN.10.2.252

Goldinger, S. D., and Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Psychol. Sci.* 21, 90–95. doi: 10.1177/0963721412436811

Granholm, E., Asarnow, R. F., Sarkin, A. J., and Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology* 33, 457–461. doi: 10.1111/j.1469-8986.1996.tb01071.x

Granholm, E., Morris, S. K., Sarkin, A. J., Asarnow, R. F., and Jeste, D. V. (1997). Pupillary responses index overload of working memory resources in schizophrenia. *J. Abnorm. Psychol.* 106, 458–467. doi: 10.1037/0021-843X.106.3.458

Hess, E. H., and Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science* 143, 1190–1192. doi: 10.1126/science.143.3611.1190

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., and Beatty, J. (1966). Pupil diameter and load on memory. *Science* 154, 1583–1585. doi: 10.1126/science.154.3756.1583

Kahneman, D., Onuska, L., and Wolman, R. E. (1968). Effects of grouping on the pupillary response in a short- term memory task. *Q. J. Exp. Psychol.* 20, 309–311. doi: 10.1080/14640746808400168

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., and Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *J. Exp. Psychol. Gen.* 133, 189–217. doi: 10.1037/0096-3445.133.2.189

Karatekin, C. (2004). Development of attentional allocation in the dual task paradigm. *Int. J. Psychophysiol.* 52, 7–21. doi: 10.1016/j.ijpsycho.2003.12.002

Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Dev. Rev.* 27, 283–348. doi: 10.1016/j.dr.2007.06.006

Karatekin, C., Marcus, D. J., and Couperus, J. W. (2007a). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. *Psychophysiology* 44, 128–144. doi: 10.1111/j.1469-8986.2006.00477.x

Karatekin, C., Marcus, D. J., and White, T. J. (2007b). Oculomotor and manual indices of incidental and intentional spatial sequence learning in middle childhood and adolescence. *J. Exp. Child Psychol.* 96, 107–130. doi: 10.1016/j.jecp.2006.05.005

Laeng, B., Sirois, S., and Gredeback, G. (2012). Pupillometry: a window to the preconscious? *Perspect. Psychol. Sci.* 7, 18–27. doi: 10.1177/1745691611427305

Laeng, B., Waterloo, K., Johnsen, S. H., Bakke, S. J., Låg, T., Simonsen, S. S., et al. (2007). The eyes remember it: oculography and pupillometry during recollection in three amnesic patients. *J. Cogn. Neurosci.* 19, 1888–1904. doi: 10.1162/jocn.2007.19.11.1888

Magimairaj, B. M., and Montgomery, J. W. (2012). Children's verbal working memory: relative importance of storage, general processing speed, and domain-general controlled attention. *Acta Psychol.* 140, 196–207. doi: 10.1016/j.actpsy.2012.05.004

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97. doi: 10.1037/h0043158

Peavler, W. (1974). Pupil size, information overload, and performance differences. *Psychophysiology* 11, 559–566. doi: 10.1111/j.1469-8986.1974.tb01114.x

Siegle, G. J., Steinhauer, S. R., Friedman, E. S., Thompson, W. S., and Thase, M. E. (2011). Remission prognosis for cognitive therapy for recurrent depression using the pupil: utility and neural correlates. *Biol. Psychiatry* 69, 726–733. doi: 10.1016/j.biopsych.2010.12.041

Simmering, V. R., and Perone, S. (2013). Working memory capacity as a dynamic process. *Front. Psychol.* 3:567. doi: 10.3389/fpsyg.2012.00567

Steinhauer, S. R., Siegle, G. J., Condray, R., and Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *Int. J. Psychophysiol.* 52, 77–86. doi: 10.1016/j.ijpsycho.2003.12.005

Unsworth, N., and Engle, R. W. (2007a). On the division of short-term and working memory: an examination of simple and complex span and their relation to higher order abilities. *Psychol. Bull.* 133, 1038–1066. doi: 10.1037/0033-2909.133.6.1038

Unsworth, N., and Engle, R. W. (2007b). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol. Rev.* 114, 104–132. doi: 10.1037/0033-295X.114.1.104

van der Meer, E., Beyer, R., Horn, J., Foth, M., Bornemann, B., Ries, J., et al. (2010). Resource allocation and fluid intelligence: insights from pupillometry. *Psychophysiology* 47, 158–169. doi: 10.1111/j.1469-8986.2009.00884.x

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children—4th Edn., Administration and Scoring Manual*. San Antonio, TX: Harcourt Assessment, Inc.

Wilhelm, B., Wilhelm, H., and Lüdtke, H. (1999). "Pupillography: principles and applications in basic and clinical research," in *Pupillography: Principles, Methods and Applications*, eds J. Kuhlmann and M. Böttcher (München: Zuckschwerdt Verlag), 1–11.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 January 2014; accepted: 25 February 2014; published online: 13 March 2014.

*Citation: Johnson EL, Miller Singley AT, Peckham AD, Johnson SL and Bunge SA (2014) Task-evoked pupillometry provides a window into the development of short-term memory capacity. Front. Psychol. **5**:218. doi: 10.3389/fpsyg.2014.00218*
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology.

Copyright © 2014 Johnson, Miller Singley, Peckham, Johnson and Bunge. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

ADVANTAGES OF PUBLISHING IN FRONTIERS

FAST PUBLICATION
Average 90 days from submission to publication

COLLABORATIVE PEER-REVIEW
Designed to be rigorous – yet also collaborative, fair and constructive

RESEARCH NETWORK
Our network increases readership for your article

OPEN ACCESS
Articles are free to read, for greatest visibility

TRANSPARENT
Editors and reviewers acknowledged by name on published articles

GLOBAL SPREAD
Six million monthly page views worldwide

COPYRIGHT TO AUTHORS
No limit to article distribution and re-use

IMPACT METRICS
Advanced metrics track your article's impact

SUPPORT
By our Swiss-based editorial team

frontiers

EPFL Innovation Park · Building I · 1015 Lausanne · Switzerland
T +41 21 510 17 00 · info@frontiersin.org · frontiersin.org