frontiers | Research Topics

When (and how) Theory of Mind is useful? Evidence from innovative assessment tools, training, and treatments strategies, volume II

Edited by
Francesca Baglio, Antonella Marchetti, Sara Isernia and Simone Shamay-Tsoory

Published in
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-7206-1
DOI 10.3389/978-2-8325-7206-1

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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: frontiersin.org/about/contact

When (and how) Theory of Mind is useful? Evidence from innovative assessment tools, training, and treatments strategies, volume II

Topic editors
Francesca Baglio — Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy
Antonella Marchetti — Catholic University of the Sacred Heart, Italy
Sara Isernia — Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy
Simone Shamay-Tsoory — University of Haifa, Israel

Citation
Baglio, F., Marchetti, A., Isernia, S., Shamay-Tsoory, S., eds. (2025). *When (and how) Theory of Mind is useful? Evidence from innovative assessment tools, training, and treatments strategies, volume II*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-7206-1

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 18 November 2025
DOI 10.3389/fpsyg.2025.1721691

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Sara Palermo,
University of Turin, Italy

*CORRESPONDENCE
Sara Isernia
sisernia@dongnocchi.it

RECEIVED 09 October 2025
ACCEPTED 04 November 2025
PUBLISHED 18 November 2025

CITATION
Baglio F, Isernia S, Shamay-Tsoory S and Marchetti A (2025) Editorial: When (and how) Theory of Mind is useful? Evidence from innovative assessment tools, training, and treatments strategies, volume II. *Front. Psychol.* 16:1721691. doi: 10.3389/fpsyg.2025.1721691

Editorial: When (and how) Theory of Mind is useful? Evidence from innovative assessment tools, training, and treatments strategies, volume II

Francesca Baglio[1], Sara Isernia[1]*, Simone Shamay-Tsoory[2] and Antonella Marchetti[3]

[1]IRCCS Fondazione Don Carlo Gnocchi ONLUS, Milan, Italy, [2]Department of Psychology, University of Haifa, Haifa, Israel, [3]Research Unit on Theory of Mind, Department of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy

KEYWORDS

theory of mind, social cognition, assessment, treatment, rehabilitation, mentalizing

Editorial on the Research Topic
When (and how) Theory of Mind is useful? Evidence from innovative assessment tools, training, and treatments strategies, volume II

Social cognition (SC) refers to several core competencies that allow individuals to successfully adapt to the interpersonal world, managing others' emotions, thoughts, and behaviors, which is essential for daily interactions and overall wellbeing (Arioli et al., 2018). Despite its importance, it has long been neglected in clinical settings, which have focused more on traditional cognitive functions. Only with the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) SC was recognized as a core neurocognitive domain (Sachdev et al., 2014). This late acknowledgment has resulted in a lack of standardized assessment tools, limiting both diagnostic accuracy and targeted intervention development (Cerami et al., 2025). This Research Topic addresses that gap, focusing on Theory of Mind (ToM), a core component of SC defined as the capacity to understand and predict behavior based on one's own and others' mental states. Building on the foundation laid by our earlier volume (Baglio and Marchetti, 2016), this Research Topic advances the field by presenting studies that develop effective assessment tools and propose innovative training strategies for therapeutic interventions. Overall the Research Topic brings together nine contributions from 44 internationally recognized authors, including seven original research articles, one perspective piece, and one opinion article. Collectively, their studies focus on developmental stages of ToM and on the research of effective tools for its assessment, as well as strategies aimed at enhancing it. Erceg et al. present a longitudinal framework of ToM measurement across the lifespan. The authors identify a critical developmental window between ages 6 and 9, which differs from adulthood, with ToM components remaining generally stable after age 9. Other studies concentrate on childhood and adolescence, stages of life in which ToM undergoes particularly significant changes. Bianco et al. address a key gap: while first-order reasoning is well-studied, second-order reasoning is less charted. Using a novel narrative-based paradigm, the authors elicited children's reasoning about characters' beliefs and desires, manipulating both valence and truth-value. Their findings reveal that second-order reasoning is more robust when associated with positive desires and true beliefs, providing important new insights into the cognitive mechanisms underpinning ToM development. Cornaggia et al. extend this research trajectory into adolescence by exploring the relationship between ToM and metalinguistic competences. Using a comprehensive assessment battery, they investigate how definitional skills and ToM interact as manifestations of broader metarepresentational abilities. Their results suggest a complex interplay: ToM performance appears to predict, albeit modestly, the ability to define ToM-related words, reinforcing long-standing evidence for a close connection between language and SC. The identification of atypical mentalizing patterns remains a central challenge in the field, particularly in the context of clinical conditions. Sharp et al. address this by introducing a novel self-report measure of hypermentalizing—where individuals draw unwarranted inferences about others' mental states. The tool is novel for being patient-reported and grounded in attachment-based theories, which allow for the consideration of different levels of everyday social relationships. The authors demonstrate the clinical utility of the tool by testing it with adolescents with personality disorders, offering a tool to identify maladaptive social cognition. Fadda et al. offer another significant advance by applying the Theory of Mind Assessment Scale (Th.o.m.a.s.) to adolescents with autism spectrum disorder. This semi-structured interview captures multiple dimensions of ToM, including Awareness (perceiving mental states in self and others), Relation (understanding causal links between mental states and behaviors), and Realization (adopting strategies to attain a desired state). Their work is enriched by the complementary review of Gabbatore et al., who critically appraise the strengths and weaknesses of widely used ToM measures. They argue that Th.o.m.a.s. is a promising tool for assessing mentalizing. At the other end of the lifespan, Sola et al. examine SC in the context of aging and neurodegenerative disorders. Using an assessment battery targeting both basic emotion recognition and affective ToM, they compare healthy older adults with individuals affected by frontotemporal dementia and Alzheimer's disease. Their findings demonstrate that frontotemporal dementia is associated with profound deficits across both domains of SC, in contrast to the comparatively milder difficulties observed in Alzheimer's disease. These results contribute to refining the clinical characterization of neurodegenerative profiles and highlight the potential diagnostic value of SC measures in differentiating neurodegenerative conditions. While assessment is crucial, translating insights into effective interventions is equally important. Two contributions address this challenge directly. Grazzani presents the PROMEHS program, which uses linguistic-conversational training activities with preschool children. Through shared story-reading, film viewing, and group discussion, the program encourages children to explore characters' inner worlds, promoting perspective-taking, emotion recognition, and the linking of emotions with behaviors. Preliminary findings suggest that this approach can foster key social competencies at an early stage, laying a foundation for lifelong social and emotional health. Birch et al. turn their attention to adolescence and adulthood, examining how interventions can address social cognitive biases as mechanisms of change. By focusing explicitly on reducing biases in social information processing, they argue that SC training can achieve more generalizable benefits, enhancing emotional wellbeing and relationships. Taken together, these contributions enrich the field of SC by advancing both theoretical and methodological perspectives. It is interesting to note that, across the studies presented here, existing measures largely remain anchored in traditional methodologies. In the future, it will be important to investigate how emerging technologies and innovative tools, such as virtual reality, can create immersive, ecologically valid scenarios that mirror real-life social interactions. Similarly, real-world (second-person) and digital phenotyping approaches may offer dynamic and fine-grained insights into brain and behavior within naturalistic contexts. Integrating these technologies into research and clinical practice represents a promising step forward—one that could shift the field from measurement toward meaningful impact, fostering interventions that are both effective and sensitive to lived experience. It is our hope that the insights offered in this Research Topic will stimulate further research, encourage cross-disciplinary dialogue, and inspire new directions in understanding SC across the lifespan.

Author contributions

FB: Project administration, Conceptualization, Writing – original draft. SI: Writing – original draft, Conceptualization, Project administration. SS-T: Project administration, Conceptualization, Writing – review & editing. AM: Project administration, Conceptualization, Writing – review & editing.

Conflict of interest

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Arioli, M., Crespi, C., and Canessa, N. (2018). Social cognition through the lens of cognitive and clinical neuroscience. *Biomed. Res. Int.* 2018:4283427. doi: 10.1155/2018/4283427

Baglio, F., and Marchetti, A. (2016). Editorial: when (and how) is theory of mind useful? Evidence from Life-Span Research. *Front. Psychol.* 7:1425. doi: 10.3389/fpsyg.2016.01425

Cerami, C., Boccardi, M., Meli, C., Panzavolta, A., Funghi, G., Festari, C., et al. (2025). Understanding barriers and optimizing socio-cognitive assessment in the diagnosis of neurocognitive disorders. *J. Neuropsychol.* 19, 603–618. doi: 10.1111/jnp.12431

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., et al. (2014). Classifying neurocognitive disorders: the DSM-5 approach. *Nat. Rev. Neurol.* 10, 634–642. doi: 10.1038/nrneurol.2014.181

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 06 August 2024
DOI 10.3389/fpsyg.2024.1439824

OPEN ACCESS

EDITED BY
Antonella Marchetti,
Catholic University of the Sacred Heart, Italy

REVIEWED BY
Ilaria Castelli,
University of Bergamo, Italy
Paola Corsano,
University of Parma, Italy

*CORRESPONDENCE
Ilaria Grazzani
✉ ilaria.grazzani@unimib.it

RECEIVED 29 May 2024
ACCEPTED 25 July 2024
PUBLISHED 06 August 2024

CITATION
Grazzani I (2024) Promoting theory of mind and emotion understanding in preschool settings: an exploratory training study. *Front. Psychol.* 15:1439824.
doi: 10.3389/fpsyg.2024.1439824

Promoting theory of mind and emotion understanding in preschool settings: an exploratory training study

Ilaria Grazzani*

Department of Human Sciences for Education "R. Massa", University of Milano-Bicocca, Milan, Italy

Introduction: This new exploratory study is part of a larger ongoing follow-up project. Its specific aim was to verify whether an innovative European Program, primarily designed to enhance children's social and emotional learning, led to gains in theory of mind and emotion understanding when implemented in preschool settings.

Methods: Thirty-four children (mean age: 56.4months; *SD*: 10.1; range: 40–70months) participated in the study. They were randomly and equally divided into a training group and a control group. The training sample completed eight linguistic-conversational activities drawn from the Program, in groups of 5 to 6 children, over 8 weeks. The activities were based on listening to stories and/or watching videos and then thinking and talking about the inner world (thoughts and emotions) of the story characters as well as the participants' own inner states. During the training phase, the children in the control group engaged in drawing or free play activities. At both the pre-test and post-test phases of the study, all the children completed a language test, a battery of theory of mind (ToM) tasks (including 'change of location' and 'unexpected content' tasks), and the Test of Emotion Comprehension which evaluates nine components of emotion understanding (EU). The validated national versions of the tests were administered in all cases.

Results: Significant differences were identified between the training and control groups. Indeed, the participants in the Program training activities, which were based on conversational exchanges between an adult and a group of children, as well as among the children themselves, outperformed the control participants on both overall theory of mind and overall emotion understanding. A more detailed analysis showed that the training group outperformed the control group in relation to both specific components of EU and the 'change of location' ToM task.

Discussion: The results of this exploratory study suggest that the Program is effective at enhancing preschoolers' social understanding and thus merits implementation in preschool settings.

KEYWORDS

theory of mind, emotion understanding, social understanding, social and emotional learning, training study, preschoolers

1 Introduction

Social understanding, which informs everyday exchanges within interpersonal relations, is the ability to interpret ourselves and others in psychological terms, and specifically as persons with inner states such as intentions, desires, emotions, beliefs, false beliefs, and other complex mental experiences. Human beings begin the long process of trying to get to grips with the social world *early* in life. Indeed, recent studies have emphasized that implicit elements of this ability can already be observed during the second year of life or even before, as in the case of infants' spontaneous helping conducts (Buttelmann et al., 2009), their displays of empathy (Bischof-Kohler, 2012), and their looking behaviors, interpreted as an indicator of their attribution of false beliefs to others (Baillargeon et al., 2018).

Indeed, social understanding encompasses both theory of mind (ToM) and emotion understanding (EU) (Tompkins et al., 2018; Grazzani and Ornaghi, 2022; Grazzani and Conte, 2024). Theory of mind has primarily been studied in relation to children's developing ability to understand that they and others possess epistemic mental states such as beliefs and false beliefs. Wellman (2014) outlined the main phases in ToM development, proposing a model whereby the child sequentially acquires an appreciation of desires, beliefs, and first- and second-order false beliefs between the ages of 2 and 8 years approximately. Harris and colleagues (Pons et al., 2004) investigated the development of emotion understanding, describing children's progressively more sophisticated comprehension of the nature, causes, and regulation of emotions between the ages of 3 and 11 years. While the literature offers well-established models of how ToM and EU generally develop as a function of age, it remains crucial to identify the factors that determine variations in children's theory of mind (ToM) and emotion understanding (EU) and shape both typical and atypical patterns of development (Montgomery et al., 2023).

Although recent studies have shown that forms of ToM and EU already feature during the pre-verbal developmental phase, language continues to be one of the most intensively investigated factors in relation to the development of social understanding (Tompkins et al., 2018). The complexity of language is well known to philosophers, linguistics, and psychologists, who – in relation to the development of ToM and EU – have focused their attention on one or more of its various dimensions, ranging from syntax to semantics and pragmatics. Indeed, metanalyses (Milligan et al., 2007) and systematic reviews (Beaudoin et al., 2020) have shown that language matters for social understanding; for example, it may facilitate the transition from implicit to explicit theory of mind and emotion understanding, given that language allows children to talk about invisible inner states and heightens their awareness of their own mental experiences.

This study is informed by the social constructivist position that adult-child conversations around mental states promote children's social understanding (Ornaghi et al., 2011; Slaughter and de Rosnay, 2017; Carpendale and Lewis, 2020). Research in this domain not only corroborates the importance of mental state language *per se* but also the role of conversation that focuses on, and directly uses, this particular kind of lexicon. Among the large number of studies that have examined this topic, many have homed in on the interaction between parent and child. Symons et al. (2006) showed that the more a mother, during her everyday conversational interactions with her child, uses mentalistic nouns, adjectives, and verbs, the better the child's later performance on false belief tasks. Adrian et al. (2005) found that the frequency and variety of the cognitive and emotional terms used by mothers during storybook reading and conversations with their preschool children were positively correlated with the latter's performance on false belief tasks. Similarly, Aram et al. (2013) showed that the cognitive and social and emotional understanding of preschool children was significantly enhanced when their parents read them stories that featured mental state lexicon and then discussed the story characters' thoughts and social interactions with them. This significant association has been borne out by more recent studies, including a German longitudinal study (Ebert et al., 2017) with preschool children in which the effect of the participants' socio-economic background was also controlled for, as well as a cross-cultural study by Taumoepeau et al. (2019) that compared samples of Australian and Iranian mothers. The latter authors only identified a significant correlation between false belief understanding and maternal mental-state talk in the group of Australian mothers, who while reading and discussing an unillustrated story with their children, used more cognitive mental state terms and referred more frequently to their own inner states.

2 Promoting social understanding in the preschool context

While there are numerous studies on parent–child interaction and conversation in family surrounding mental states and inner experience, less research has been carried out in extrafamilial contexts, as shown in a meta-analysis by Hofmann et al. (2016). The study of conversational interaction about mental states between educators/teachers and children at nursery or kindergarten can concern both the adult-individual child interaction and the interaction between an adult and a small group of children. The empirical focus of the present work is on this second possibility.

Conversation between an adult and a small group of children can in principle foster the development of abilities that are key for social understanding. In the course of group conversations, children are encouraged to listen to the utterances of others (e.g., statements, comments, questions, and answers), to put themselves in the shoes of others, and to compare their own point of view with that of others. This kind of activity can enhance the perspective-taking ability that is crucial to understanding the social world and to recognizing that – with respect to oneself – another person may hold different or similar intentions, perceptions, thoughts, beliefs, emotions, feelings, needs, motives, and information (Carpendale and Lewis, 2020).

To date, few studies have explored how children's ToM and EU may be enhanced via programs based on activities with small groups of children in preschool educational settings. We define 'programs' here as structured interventions with accompanying guidelines that can help adults in educational contexts (e.g., educators, teachers, education specialists, psychologists, and so on) to conduct targeted activities with a view to improving children's social understanding. In a pioneering study focused on the Italian cultural context and preschool children, Ornaghi et al. (2011) implemented a two-month intervention in kindergartens, during which 3- and 4-year-olds were read stories enriched with mental state terms. After listening to a story, the participants took part in guided language games and conversations aimed at stimulating their use of a variety of mental state terms. As compared to the control group, the intervention group displayed significant gains in their theory-of-mind and metacognitive vocabulary comprehension. This intervention informed the development of a structured program based on language games around mental terms, including 'to think', 'to be afraid', 'to decide', 'to remember' and so on (Grazzani and Ornaghi, 2022). In addition, Grazzani and Ornaghi (2011) and Ornaghi et al. (2015) found similar results with children aged 3 to 5 years who participated in emotion language games and subsequently displayed significant gains in their emotion understanding and use of emotional lexicon. However, to date, no intervention program has been associated with either of these two studies. In a study by Mori and Cigala (2019), small groups of preschool children were trained in perspective taking. The significant outcomes obtained prompted the development of a program for schools that targets multiple dimensions of perspective taking, from the perceptive to the cognitive dimension (Mori and Cigala, 2021). The program offers a wide variety of structured activities involving movement, drawing, and dramatization. The promotion of theory of mind in preschool age is in continuity with that enhanced in primary school context. Bianco et al. (2019), for instance, showed that conversation about mental states allowed children between 7 and 8 years of age to outperform in ToM skills children belonging to the control group. Similar findings were described in Bianco et al. (2021) with a large sample of students who showed a significant training effect on advanced-ToM and metacognition abilities. A validated program for improving theory of mind in primary school children is presented in the volume by Lecce and Bianco (2018) which includes precise guidelines for teachers.

The PROMEHS program/curriculum, which is partly based on conversational activities, has recently been validated as an effective means of enhancing the social and emotional learning (SEL) of preschool, primary, and secondary school students (Cefai et al., 2022). The Program targets the five main components of SEL (Durlak et al., 2015; CASEL-Collaborative on Academic Social Emotional Learning, 2020; Cavioni et al., 2020a): self-awareness (the ability to identify one's own thoughts, emotions, and needs), self-management (the ability to regulate one's own thoughts, emotions, and behaviors), social awareness (the ability to understand another person's point of view, thoughts, and emotions), relationship skills (the ability to establish and maintain positive social relationships) and responsible decision making (the ability to make informed decisions). Each of the five components corresponds to a specific set of competences or sub-skills well described by the CASEL's framework. For example, self-awareness includes awareness of one's own thoughts and emotions and how these relate to one's own actions; social awareness concerns the perspective-taking process and empathic behavior; relationship skills encompass the ability to coordinate with others, which presupposes the capacity to recognize their needs, intentions, thoughts, and so on (Mahoney et al., 2021). Figure 1 illustrates the relationship between the five SEL components on the one hand, and the activities implemented during the training which targeted these subskills on the other hand. It can be seen that these sub-skills evoke the above definition of social understanding in terms of theory of mind and emotion comprehension, a definition that underlines the awareness of one's own thoughts and emotions as well as the distinction between one's own and others' internal states (Wellman, 2014).

During the original experimental evaluation of the PROMEHS Program (Cefai et al., 2022), its effectiveness with preschoolers was only verified indirectly by asking the children's teachers to complete the SSIS-SEL Brief Scales by Elliott et al. (2020). As noted in Conte et al. (2023), this instrument now validated in Italian (Cavioni et al., 2023) comprises 20 items to be rated on a 4-point Likert scale ranging from 1, "not true" to 4, "very true." It was used to indirectly assess participants' SEL competences before and after the experimental intervention. As such, it did not provide direct evidence of the Program's impact on children's skills.

3 The present study

In light of the background outlined above, we conducted a new training study with two groups of participants: an experimental group (or training group) and a control group. The study had two main aims:

(1) Given that the effectiveness of PROMEHS Program had not previously been tested using direct measures, in this study we set out to implement part of the program via a targeted training intervention with a new, smaller sample of children, and to assess the impact of the training by administering direct measures;
(2) Given that – as noted above – the components of SEL overlap with those of social understanding, in this study we set out to evaluate the effectiveness of the PROMEHS Program by administering direct measures of the participating children's theory of mind and emotion understanding. Differences between the training and control groups at post-test would corroborate the hypothesis that the Program activities under study, which were specifically devised to promote SEL, may usefully be included in interventions targeting preschoolers' social understanding as well, that is to say, their theory of mind and emotion understanding abilities.

Based on the results of earlier training studies conducted using a conversational approach (for a review: Tompkins et al., 2018), we expected to find significant differences between the experimental group (or training group) and the control group following the intervention.

4 Method

4.1 Participants

The participants in this study were 34 children (16 girls; mean age at the beginning of the study = 56.4 months, range = 40–70 months; $SD = 10.1$). Evenly divided into two groups (a training and a control group), the children were enrolled at two nursery schools in an urban area of a Northern Italian region. The head teachers agreed for the schools to participate in the study because of the potential benefits for their preschools and teachers. All the children were native Italian speakers. The inclusion criteria for this study were: (1) the children were required to be aged between 40 and 72 months, and (2) the parents were required to provide informed consent for their children's participation in the study. The parents, who were of medium socioeconomic status, attended a presentation of the study, and then signed the informed consent forms for their children. This approach to obtaining informed consent was in line with the Declaration of Helsinki principles. Participants were free to withdraw from the study at any time. The study was conducted in conformity with the recommendations of the University of Milano-Bicocca Ethics Committee.

4.2 Measures

The research instruments selected were validated measures and appropriate to the age of the participants. We followed Denham's

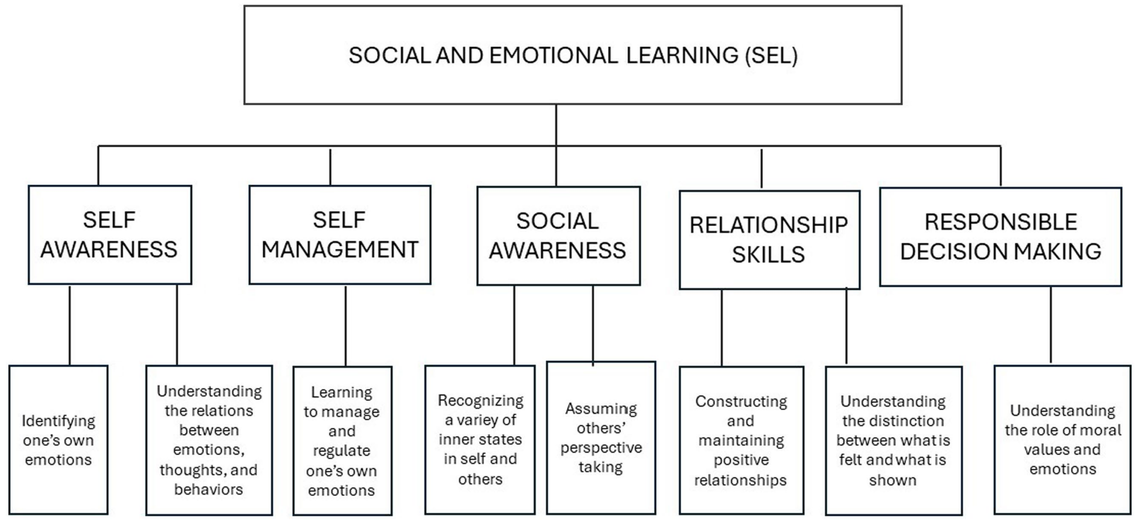


FIGURE 1
SEL components ant the related trained activities.

recommendations (Denham et al., 2010) who suggested evaluating the effectiveness of evidence-based SEL programs through tasks directly administered to children. Both before and after the training phase, the children completed a battery of tests, in counterbalanced order, in a quiet and familiar room at their nursery school. More specifically, the following tests were administered by a member of the research team who had spent familiarization time getting to know the children beforehand.

4.2.1 Peabody picture vocabulary test

The PPVT (Dunn and Dunn 2007) was administered both to verify that the two groups were starting out with equivalent levels of linguistic ability and to verify that none of the participants displayed strongly atypical language development. We use the Italian standardized version of the test (Stella Pizzoli and Tressoldi, 2000). It evaluates the receptive vocabulary of children between 3 and 12 years and consists of 180 cards, each containing four numbered illustrations among which the child is asked to indicate the one that corresponds to the word called out by the examiner. Scoring was carried out following the standard procedure, with 1 point assigned for each correct answer and 0 for each wrong answer. The reliability coefficient was $\alpha = 0.70$.

4.2.2 Test of emotion comprehension

The TEC devised by Pons and Harris (2000) assesses emotion comprehension in 3- to 11-year-olds. It encompasses nine components (Pons et al., 2004), namely: the recognition of facial expressions of emotions; the understanding of, respectively, the impact of situational causes on emotions, the role of desires in emotions, the role of beliefs in emotions, the impact of memory on emotions, the distinction between outwardly expressed and privately felt emotions, and the effect of morality on emotions; the awareness that emotions may be regulated using cognitive control strategies; and, finally, an appreciation of concurrent mixed feelings. In the current study, we deployed the standardized Italian version (Albanese and Molina, 2008). The TEC assesses emotion understanding by presenting vignettes in which a gender-matched protagonist encounters simple to complex situations eliciting different emotional responses. After each vignette, the child is asked to indicate how the protagonist feels, by choosing among four illustrations of faces representing different emotional states. For each group of items testing an individual component, a score ranging from 0 to 1 is awarded. These scores are then summed to obtain a Total TEC score, which ranges from 0 to 9. The scoring system was defined and applied in strict accordance with the guidelines of Pons and Harris (2000) and subsequent recommendations by Cavioni et al. (2020a,b). The reliability coefficient was $\alpha = 0.73$.

4.2.3 ToM battery

This measure consisted of two first-order false-belief tests: (a) a 'false-belief location change task' consisting in the Italian adaptation of the classic "Sally and Ann" story (Liverta Sempio et al., 2005); (b) a false-belief unexpected content task consisting in the Italian adaptation by Liverta Sempio et al. (2005). For each task, the children were awarded scores of 1 for the correct answer and 0 for the wrong answer. Scores for the battery were summed to yield a possible maximum total score of 2. The reliability coefficient was $\alpha = 0.80$.

4.3 The program and training procedure

Following the pre-test phase, the training phase was implemented. The children in the training group participated in eight activities lasting approximately 45 min each, offered once a week over a period of eight weeks. The training activities were selected from the PROMEHS Program (for further information, see Grazzani et al., 2020) as a function of the children's age and of the aim of the study, which was to foster competences relating to the SEL components. Figure 1 shows how the components of SEL are related to the respective target activities and competences.

The training was conducted by a member of the research team with weekly supervision by the project leader. The trainer spent a week developing a relationship with the children before initiating the intervention. Each training session involved a small group of children and was divided into two parts. First, the trainer read aloud to the children a story drawn from the PROMEHS Program (e.g., 'The three little pigs'). Second, the children took part in conversational activities guided by the trainer, who followed the guidelines provided in the manual of the Program. The children were prompted to actively participate in the conversation by answering questions and recounting and sharing their own experiences, emotions, and thoughts. During the conversational activities, the trainer emphasized the subjective nature of mental states and encouraged the children both to discuss their own perspectives and to adopt the perspectives of others. The children in the control group simply listened to the stories and then participating in unstructured drawing activities. All the teachers who participated in the study agreed to be trained after the post-test phase. Supplementary Appendix A1 outlines a sample of activities from the Program whose aim is to foster the ability to "identify and name basic emotions," "learn to regulate one's own emotions," "assume others' perspective taking," "construct and maintain positive relationship" and "understand the role of moral values and emotions."

TABLE 1 Means and standard deviations of all variables both before and after the training phase.

	MEAN (pre and post)	SD (pre and post)
Age in months	56.48–58.84	10.158–10.080
Languge ability (Peabody)	53.45–60.32	23.490–25.718
Emotion understanding (TEC)		
TEC - Component 1	0.94–0.97	0.250–0.180
TEC - Component 2	0.81–0.74	0.402–0.445
TEC - Component 3	0.52–0.71	0.508–0.461
TEC - Component 4	0.52–0.65	0.508–0.486
TEC - Component 5	0.35–0.71	0.486–0.461
TEC - Component 6	0.29–0.42	0.461–0.502
TEC - Component 7	0.48–0.55	0.508–0.506
TEC - Component 8	0.06–0.23	0.250–0.425
TEC - Component 9	0.52–0.61	0.508–0.495
TEC - Total score	4.48–5.58	1.730–2.277
Theory of mind (Battery of tasks)		
ToM (Change of location)	0.39–0.58	0.495–0.502
ToM (Unexpected content)	0.32–0.45	0.475–0.506
ToM - Total score	0.71–1.03	0.824–0.875

5 Results

All the data analyses in this exploratory study were conducted using the software IBM SPSS Version 29. Before analyzing the efficacy of the training, standard data-cleaning procedures were conducted. No missing values were detected. Anticipating a medium effect size based on prior research in this domain, we set the desired *power level a priori* at 0.70. The analysis indicated that a minimum of 16 participants per group condition would be required to obtain 95% statistical power in detecting the expected effect. The data were matched by code to combine the pre- and post-test scores; all the children received scores for both tests.

The results section comprises two subsections outlining descriptive statistics for all study measures and the impact of the training on children's theory of mind and emotion comprehension, respectively. Please recall that, given the linguistic-conversational nature of the training, the language measure was administered amongst other reasons with a view to identifying any atypical patterns of language development.

5.1 Descriptive statistics

Descriptive statistics ($n = 34$) at Time 1 and Time 2 for the variables under study are reported in Table 1. These include the means and standard deviations, both before and after the training phase, of the following: age in months, language ability as assessed by the Peabody Test, emotion understanding as evaluated by the TEC (including each of the nine components mentioned in the Measures section), and theory of mind as assessed via the ToM battery with the change of location and unexpected content tasks. Correlations were calculated using the pre-test data. Age in months was positively and significantly associated with language ($r = 0.70$, $p < 0.001$), Total TEC ($r = 0.63$, $p < 0.001$) and Total ToM ($r = 0.50$, $p = 0.004$); language was significantly associated with both Total TEC ($r = 0.64$, $p < 0.001$) and Total ToM ($r = 0.41$, $p = 0.02$); there was a significant association between EU and ToM, and more specifically between Component 9 of EU (an appreciation of concurrent mixed feelings) and Total ToM ($r = 0.37$, $= p = 0.04$).

5.2 The effectiveness of the training

At pre-test, the training and control groups did not differ significantly in relation to any of the dependent variables, namely language ability ($p = 0.537$), emotion understanding ($p = 0.333$), and theory of mind ($p = 0.567$). No differences emerged as a function of gender; therefore, this variable was not included in the subsequent analyses.

To test the impact of the training activities, a repeated measures multivariate analysis of variance (Manova) was carried out. The independent variables were Time (pre-test and post-test) and Group Condition (training vs. control); Time was a within-participant variable whereas Group Condition was a between-participant variable. The dependent variables at the two time points were language ability, emotion understanding (Total TEC), and theory of mind (Total ToM). Effect sizes were calculated using partial eta-squared (η_p^2) values.

Table 2 presents the means and standard deviations at Time 1 (pre-test) and Time 2 (post-test) as a function of Group Condition (training vs. control group).

Concerning the effect of the training on emotion understanding (Total TEC), there was a significant Time x Group Condition interaction, Wilks's $\lambda = 0.80$, $F = 6{,}99$, $p = 0.01$, $\eta_p^2 = 0.194$. The

TABLE 2 Pre- and post-test means and standard deviations for all variables by group condition.

	Training group		Control group	
	Pre-test	Post-test	Pre-test	Post-test
Age in months	57.67 (8.9)	60.07 (8.8)	55.38 (11.9)	57.69 (11.2)
Languge ability (Peabody)	56.20 (21.2)	64.47 (25.3)	50.88 (25.8)	56.44 (26.3)
EU (TEC)				
TEC - Component 1	0.93 (0.25)	1.00 (0.0)	0.94 (0.25)	0.94 (0.25)
TEC - Component 2	0.87 (0.35)	0.93 (0.25)	0.75 (0.44)	0.56 (0.51)
TEC - Component 3	0.53 (0.51)	0.93 (0.25)	0.50 (0.51)	0.50 (0.51)
TEC - Component 4	0.47 (0.51)	0.67 (0.48)	0.56 (0.51)	0.63 (0.50)
TEC - Component 5	0.47 (0.51)	0.80 (0.41)	0.25 (0.44)	0.63 (0.50)
TEC - Component 6	0.33 (0.48)	0.53 (0.51)	0.25 (0.44)	0.31 (0.47)
TEC - Component 7	0.47 (0.51)	0.80 (0.41)	0.50 (0.51)	0.31 (0.47)
TEC - Component 8	0.7 (0.25)	0.27 (0.45)	0.6 (0.25)	0.19 (0.40)
TEC - Component 9	0.67 (0.48)	0.80 (0.41)	0.38 (0.50)	0.44 (0.51)
TEC – Total Score	4.80 (1.6)	6.73 (1.4)	4.19 (1.7)	4.50 (2.4)
ToM (Battery of tasks)				
ToM 1 Change of loc.	0.33 (0.48)	0.80 (0.41)	0.44 (0.51)	0.38 (0.50)
ToM 2 Unexpected content	0.47 (0.51)	0.60 (0.50)	0.19 (0.40)	0.31 (0.47)
ToM – Total Score	0.80 (0.86)	1.40 (0.73)	0.63 (0.80)	0.69 (0.87)

Standard deviations are in parentheses.

univariate test showed that the training group outperformed the control group, with the former displaying a significantly greater gain in their global emotion understanding, $F = 9{,}57$, $p = 0.004$, $\eta_p^2 = 0.248$. from pre- to post-test.

Concerning the impact of the training on ToM (Total score), there was a significant Time x Group Condition interaction, Wilks's $\lambda = 0.74$, $F = 10.04$, $p = 0.004$, $\eta_p^2 = 0.257$. The univariate test showed that the training group outperformed the control group, in that the former displayed a significantly greater gain from pre- to post-test in their global theory of mind, $F = 5{,}98$, $p = 0.02$, $\eta_p^2 = 0.171$. The interaction Time x Group Condition was not significant for language ability, Wilks's $\lambda = 0.98$, $F = 0{,}33$, $p = 0.57$, $\eta_p^2 = 0.011$.

Analyses for the Time factor revealed that the control group displayed a significant pre- to post-test increase in language ability, $t\ (15) = 8{,}58$, $p < 0.001$, TEC Total score, $t\ (15) = 7{,}432$, $p < 0.001$, and ToM Total score, $t\ (15) = 3{,}149$, $p = 0.007$. Similarly, the training group displayed a significant pre- to post-test increase in language ability, $t\ (15) = 9{,}866$, $p < 0.001$, TEC Total score, $t(15) = 18{,}140$, $p < 0.001$, and ToM Total score, $t\ (15) = 9{,}866$, $p < 0.001$.

Figures 2, 3 illustrate the significantly greater improvement of the training group – as compared to the control group – from pre-test (Time 1) to post-test (Time 2) in global EU as well as in global ToM.

When we analyzed the impact of the training on specific components of the TEC, we identified significant effects on Component 2, understanding the impact of situational causes on emotions ($F = 6{,}33$, df = 1, $p = 0.01$), Component 3, understanding the role of desire in emotions ($F = 8{,}54$, df = 1, $p = 0.007$), Component 7, the effect of morality on emotions ($F = 9{,}14$, df = 1, $p = 0.005$), and Component 9, the appreciation of concurrent mixed feelings ($F = 4{,}65$, df = 1, $p = 0.03$). Finally, analysis of the impact of the training on the individual tests in the Theory of Mind Battery showed that the training group improved significantly more on the 'change of location' test ($F = 6{,}59$, df = 1, $p = 0.01$) than did the control group.

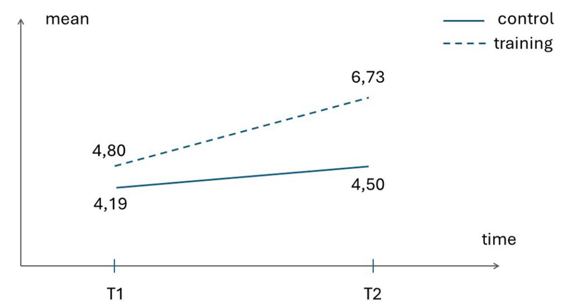


FIGURE 2
Control and Training groups' scores on emotion understanding at pre-test (Tl) and post-test (T2).

6 Discussion

This study built on the previous research of Conte et al. (2023) which was designed to verify the effectiveness of the PROMEHS program in promoting preschoolers' social and emotional learning. In the study by Conte et al. (2023), however, the effectiveness of the program was only tested using indirect measures, namely teacher-report questionnaires (Elliott et al., 2020). Identifying appropriate direct measures and administering them to a fresh sample of children

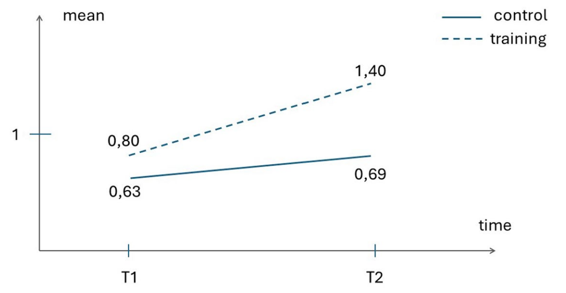


FIGURE 3
Control and Training groups' scores on theory of mind at pre-test (Tl) and post-test (T2).

remained an open challenge. In the current study, one objective was thus to evaluate a part of the PROMEHS Program for preschoolers via direct measures, in line with the recommendations of Denham et al. (2010).

Furthermore, given the link between the components of SEL and those of social understanding as illustrated in Figure 1, we selected direct measures of theory of mind and emotion understanding to investigate the effectiveness of the program. The proposal to evaluate the impact of a SEL curriculum also through theory of mind measures is innovative as compared to the existing findings. We expected that we would find differences in post-test performance between the children who had participated in the experimental (training group) and those in the control group who had only listened to stories, watched videos, or engaged in unstructured drawing activities without participating in the conversational activities. The data analysis shows that at Time 2 (post-test), the mean test performance of both groups of children had improved relative to Time 1. However, further analyses confirmed our research hypothesis, showing that the children in the training group displayed significantly greater gains in both theory of mind and emotion understanding than did the children in the control group. This effect was more powerful with respect to emotion understanding (effect size: 0.248) than with respect to theory of mind (effect size: 0.171). Both effect sizes were modest, yet in line with those found in similar past studies (e.g., Ornaghi et al., 2015). Plausibly, larger effect sizes might be found by increasing the size of the research sample.

PROMEHS' relatively greater impact on EU may be explained by the characteristics of the Program itself, which includes a high proportion of activities that engage the emotional sphere and specifically target skills such as the recognition of emotions and their causes, and the ability to regulate ongoing emotional experience. The data analysis showed that particularly significant gains were displayed in components of EU that directly featured in the training activities, such as understanding the role of desires in emotions as well as the effect of morality on feelings.

It may be concluded that regularly conducting conversational activities with small groups of children over an approximately two-month period favored the development of the skills required to perform well in the tests, such as the ability to adopt the perspective of others and compare it to one's own, as well as the ability to identify the characteristics of different emotions and to grasp their relationship with manifest behaviors. The training intervention was designed to maximize conversational exchanges between the adult and the children as well as among the children themselves. In the course of the training, this objective was increasingly more fully attained, as the children became more and more familiar with this kind of activity and increasingly more adept at actively contributing to the conversations. This implied being able to progressively talk more about themselves and others, rather than about the characters in the stories, and to contribute in an original and non-imitative way to the conversational exchanges. To this regard, consider this brief language exchange among children involved in a conversation about anger regulation, in which they manage the exchange without the need of teacher intervention. *Teacher*: What can we do to make the anger go away? *Paolo*: We can drink some water. *Anna*: eh …but that does not make the anger go away. *Paolo*: at best [it works] when you cry and are very angry. *Giulia*: you can at least breath *Paolo*: you can stay cool.

Listening to the stories and to the stimulus questions and input of the trainer prompted the children to actively join in conversations about inner states and to deploy a variety of terms from the psychological lexicon, including emotional, cognitive and volitive terms. Through language and conversation, the children were becoming more acutely aware of thoughts and emotions as causes of actions and also, vice versa, actions as causes of emotions and thoughts. Moreover, the conversational activities encouraged the children to engage in processes of metacognitive reflection about their own and others' internal states and to recognize the difference between the external causes of emotions and internal causes such as memories and desires. The children were also encouraged to take the perspective of the story characters, and to explicitly state what they would have done and how they would have felt had they been in these characters' shoes. They were also prompted to explicitly discuss how to deal with complex situations such as feeling too sad or extremely angry, or when others display strong negative emotions. The children were also invited to think about good and polite behaviors (their own and others) that made them feel good (such as cooperating, helping, sharing toys, consoling) and to contrast these with behaviors that made them feel bad (e.g., fighting, hitting, snatching toys).

Finally, it should be noted that this program was initially implemented to foster social and emotional learning hence its efficacy in enhancing children's social understanding could not be taken for granted. Furthermore, the competences associated with the five components of SEL targeted by the Program are not exclusively related to social understanding, given that they also include prosocial conduct, decision making, and problem solving. Hence, a program devised to promote SEL is not entirely comparable to a program for the development of social understanding understood as ToM and EU. The key findings of the present study are therefore that implementing the PROMEHS SEL Program with preschool children is indeed linked with gains in social understanding as well, while the direct measures administered effectively captured these improvements.

7 Limitations, strengths, and implications

The limitations of this study should be noted. First, this was an exploratory study with a small number of participants. More studies and more data are required to corroborate these interesting preliminary findings on the effectiveness of this program in enhancing preschoolers' theory of mind and emotion understanding.

Second, broader studies are required to take into account other variables that have not been considered here, such as the socio-economic status of the participants, possible differences in the effectiveness of the program as a function of age, alternative control group activities (for example a group that does drama instead of conversation activities), and longitudinal follow-up data. In addition, the size of this sample did not allow us to delve into the role of other factors such as the presence of siblings, which deserve attention in future research.

Third, drawing on the most recent theorizing surrounding adult-child conversations about mental states (Tompkins et al., 2022; Farrell et al., 2023), more detailed investigations are required to maximize the impact of programs such as that implemented in the present study. In this regard, the nature of the children's contributions to the conversational exchanges (for example, appropriate vs. off-topic comments) on the one hand, and the adult's linguistic style (for instance, tending to use open-ended questions versus closed questions) deserve more in-depth scrutiny. This knowledge could then be applied when training education practitioners to implement the program. Importantly, most studies that have drawn on the social constructivist perspective informing our work here have focused on the interaction between one adult (e.g., a parent) and one child rather than between an education practitioner and a group of children.

Despite these limitations, the findings of this study suggest that the PROMEHS Program has a significant impact in terms of effectively fostering not only children's understanding of emotions, as might be hoped of a program that intentionally targets social and *emotional* learning, but also their theory of mind abilities. We therefore believe that this program, which is already available in numerous European languages, can produce positive effects when used in preschool settings by appropriately trained practitioners. In these educational contexts, access to structured programs with materials and guidelines can greatly facilitate the implementation of targeted educational activities with children here. In addition, this kind of program could be particularly useful for those children who have poor socio-emotional skills or show a delay in linguistic development also due to an immigration background, in line with the results found by Conte et al. (2023) with children with an atypical profile. Listening to stories, the encouragement to speak through stimulating questions, listening to the classmates' comments and responses create, in fact, a favorable context for the development of cognitive, linguistic and socio-emotional skills. In sum, this study contributes to the existing body of applied research that is informed by the concept that social understanding may be taught in kindergarten (Grazzani and Brockmeier, 2019), a privileged setting for exploring how children come to understand their social world and for 'validating' the best educational practices for enhancing this understanding.

Data availability statement

The raw data supporting the conclusions of this article and the material of the Program will be made available by the author as appropriate.

Ethics statement

The study was conducted in conformity with the recommendations of the University of Milano-Bicocca Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

IG: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was funded by a grant from the University of Milano-Bicocca [ATE-FA 2020] awarded to the IG.

Acknowledgments

We would like to thank the head teacher and the teachers who facilitated the data collection in the preschools, as well as the children who participated in the study, and their parents who authorized their participation. We would like to thank the Lab-PSE team for their help with collecting the data and Clare O'Sullivan for the linguistic revision of this article.

Conflict of interest

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1439824/full#supplementary-material

References

Adrian, J. E., Clemente, R. A., Villanueva, R., and Rieffe, C. (2005). Parent-child picture-book reading, mothers' mental state language, and children's theory of mind. *J. Child Lang.* 32, 673–686. doi: 10.1017/S0305000905006963

Albanese, O., and Molina, P. (2008). Lo sviluppo della comprensione delle emozioni e la sua valutazione. Unicopli: Milano.

Aram, D., Fine, Y., and Ziv, M. (2013). Enhancing parent-child shared book reading interactions: promoting references to the book's plot and socio-cognitive themes. *Early Child Res. Q.* 28, 111–122. doi: 10.1016/j.ecresq.2012.03.005

Baillargeon, R., Buttelmann, D., and Southgate, V. (2018). Invited commentary: interpreting failed replications of early false-belief findings: methodological and theoretical considerations. *Cogn. Dev.* 46, 112–124. doi: 10.1016/j.cogdev.2018.06.001

Beaudoin, C., Leblanc, E., Gagner, C., and Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Front. Psychol.* 10:2905. doi: 10.3389/fpsyg.2019.02905

Bianco, F., Lombardi, E., Lecce, S., Marchetti, A., Massaro, D., Valle, A., et al. (2021). Supporting children's second-order recursive thinking and advanced ToM abilities: a training study. *J. Cogn. Dev.* 22, 561–584. doi: 10.1080/15248372.2021.1901712

Bianco, F., Lombardi, E., Massaro, D., Castelli, I., Valle, A., Marchetti, A., et al. (2019). Enhancing advanced theory of mind skills in primary school: a training study with 7- to 8-year-olds. *Infant Child Dev.* 28:e2155. doi: 10.1002/icd.2155

Bischof-Kohler, D. (2012). Empathy and self-recognition in phylogenetic and ontogenetic perspective. *Emotion* 4, 40–48. doi: 10.1177/1754073911421377

Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-olds understand false-belief in an unexpected-contents task. *J. Exp. Child Psychol.* 119, 120–126. doi: 10.1016/j.jecp.2013.10.002

Carpendale, J. J., and Lewis, C. (2020). What makes us human. How mind develops through social interactions. London: Routledge.

CASEL-Collaborative on Academic Social Emotional Learning. (2020). SEL competency framework. CASEL. https://casel. org/sel-framework/

Cavioni, V., Conte, E., Grazzani, I., Ornaghi, V., Cefai, C., Anthony, C., et al. (2023). Validation of Italian students' self-ratings on the SSIS SEL brief scales. *Front. Psychol.* 14:1229653. doi: 10.3389/fpsyg.2023.1229653

Cavioni, V., Grazzani, I., and Ornaghi, V. (2020a). Mental health promotion in schools: a comprehensive theoretical framework. *Int. J. Emot. Educ.* 12, 65–82. doi: 10.3389/fpsyg.2021.720628

Cavioni, V., Grazzani, I., Ornaghi, V., Pepe, A., and Pons, F. (2020b). Assessing the factor structure and measurement invariance of the test of emotion comprehension (TEC): a large cross-sectional study with children aged 3-10 years. *J. Cogn. Dev.* 21, 406–424. doi: 10.1080/15248372.2020.1741365

Cefai, C., Camilleri, L., Bartolo, P., Grazzani, I., Cavioni, V., Conte, E., et al. (2022). The effectiveness of a school-based, universal mental health programme in six European countries. *Front. Psychol.* 13:925614. doi: 10.3389/fpsyg.2022.925614

Conte, E., Cavioni, V., Ornaghi, V., Agliati, A., Gandellini, S., Santos, M. F., et al. (2023). Supporting preschoolers' mental health and academic learning through the PROMEHS program: a training study. *Children* 10:1070. doi: 10.3390/children10061070

Denham, S. A., Ji, P., and Hamre, B. (2010). Compendium of preschool through elementary school social-emotional learning and associated assessment measures. Chicago: University of Illinois.

Dunn, L. M., and Dunn, D. M. (2007). Peabody picture vocabulary test, Amsterdam. Pearson Edn. Chicago.

Durlak, J. A., Domitrovich, C. E., Weissberg, R. P., and Gullotta, T. P. (Eds.) (2015). Handbook of social and emotional learning. Research and practice. New York. Amsterdam. Guilford Press.

Ebert, S., Peterson, C., Slaughter, V., and Weinert, S. (2017). Links among parents' mental state language, family socioeconomic status, and preschoolers' theory of mind development. *Cogn. Dev.* 44, 32–48. doi: 10.1016/j.cogdev.2017.08.005

Elliott, S. N., DiPerna, J. C., Anthony, C. J., Lei, P., and Gresham, F. M. (2020). Social skills improvement system, social emotional learning (SSIS-SEL) brief scales. Scottsdale, AZ: SAIL CoLab.

Farrell, C., Slaughter, V., Thai, M., and Mulvihill, A. (2023). How we talk to kids: adults prefer different forms of language for children based on gender expression. *Sex Roles* 89, 119–134. doi: 10.1007/s11199-023-01393-7

Grazzani, I., Agliati, A., Conte, E., Cavioni, V., Gandellini, S., Ornaghi, V., et al. (2020). PROMEHS. Il curriculum per promuovere la salute mentale a scuola. Milano: Mimesis.

Grazzani, I., and Brockmeier, J. (2019). Language games and social cognition: Revisiting Bruner. *Integr Psych Behav*, 53, 602–610. doi: 10.1007/s12124-019-09489-0

Grazzani, I., and Conte, E. (2024). Lo sviluppo della comprensione sociale. Una rassegna narrativa focalizzata sul linguaggio nella fascia 0-6 anni (The development of social cognition in the 0-6 years period. A narrative review focused on the role of language). *Psicol. Clin. Svilupp.* 18, 29–58. doi: 10.1449/112819

Grazzani, I., and Ornaghi, V. (2011). Emotional state talk and emotion understanding A training study with preschool children. *J. Child Lang.* 38, 1124–1139. doi: 10.1017/S0305000910000772

Grazzani, I., and Ornaghi, V. (2022). *Children's development of social understanding*. Parma: Junior.

Hofmann, S. G., Doan, S. N., Sprung, M., Wilson, A., Ebesutani, C., Andrews, L. A., et al. (2016). Training children's theory-of-mind: a meta-analysis of controlled studies. *Cogn.* 150, 200–212. doi: 10.1016/j.cognition.2016.01.006

Lecce, S., and Bianco, F. (2018). *Mettiti nei miei panni. Programma per promuovere l'empatia e le relazioni con la teoria della mente*. Trento: Erickson.

Liverta Sempio, O., Marchetti, A., Castelli, I., Lecciso, F., and Pezzotta, C. (2005). *Mentalizzazione e competenza sociale*. Milano: FrancoAngeli.

Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R. J., Niemi, K., et al. (2021). Systemic social and emotional learning: promoting educational success for all preschool to high school students. *Am. Psychol.* 76, 1128–1142. doi: 10.1037/amp0000701

Milligan, K., Astington, J. W., and Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Dev.* 78, 622–646. doi: 10.1111/j.1467-8624.2007.01018.x

Montgomery, D. E., Tompkins, V., and Benigno, J. P. (2023). "Theory of mind" in Handbook of clinical child psychology: Integrating theory and research into practice. ed. J. L. Matson (New York: Springer), 269–287.

Mori, A., and Cigala, A. (2019). 'Putting oneself in someone's else shoes during childhood: how to learn it' training for preschool age children. *Br. J. Edu. Psychol.* 89, 750–766. doi: 10.1111/bjep.12255

Mori, A., and Cigala, A. (2021). Cambiamo la prospettiva [tr. Let's change the perspective]. Trento: Erickson.

Ornaghi, V., Brockmeier, J., and Grazzani, I. (2011). The role of language games in children's understanding of mental states: a training study. *J. Cogn. Dev.* 12, 239–259. doi: 10.1080/15248372.2011.563487

Ornaghi, V., Grazzani, I., Cherubin, E., Conte, E., and Piralli, F. (2015). "Let's talk about the emotions". The effect of conversational training on preschoolers' emotion comprehension and prosocial orientation. *Soc. Dev.* 24, 166–183. doi: 10.1111/sode.12091

Pons, F., and Harris, P. L. (2000). TEC (Test of Emotion Comprehension). Oxford: Oxford University. [this is unpublished and it is commonly cited this way].

Pons, F., Harris, P. L., and de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: developmental periods and hierarchical organization. *Eur. J. Dev. Psychol.* 1, 127–152. doi: 10.1080/17405620344000022

Slaughter, V., and de Rosnay, M. (2017). Theory of mind development in context. New York: Routledge.

Stella Pizzoli, C., and Tressoldi, P. (2000). Peabody. Test di Vocabolario recettivo. Torino: Omega.

Symons, D. K., Fossum, K., and Collins, T. B. K. (2006). A longitudinal study of belief and desire state discourse during mother-child play and later false belief understanding. *Soc. Dev.* 15, 676–692. doi: 10.1111/j.1467-9507.2006.00364.x

Taumoepeau, M., Sadeghi, S., and Nobilo, A. (2019). Cross-cultural differences in children's theory of mind in Iran and new Zeland: the role of caregiver mental state talk. *Cog. Dev.* 51, 32–45. doi: 10.1016/j.cogdev.2019.05.004

Tompkins, V., Benigno, J. P., Kiger Lee, B., and Wright, B. M. (2018). The relation between parents' mental state talk and children's social understanding: a meta-analysis. *Soc. Dev.* 27, 223–246. doi: 10.1111/sode.12280

Tompkins, V., Montgomery, D. E., and Blosser, M. K. (2022). Mother-child talk about mental states: the what, who, and how of conversations about the mind. *Soc. Dev.* 31, 281–302. doi: 10.1111/sode.12551

Wellman, H. M. (2014). Making minds. How theory of mind develops. New York: Oxford University Press.

TYPE Original Research
PUBLISHED 12 August 2024
DOI 10.3389/fpsyg.2024.1456432

OPEN ACCESS

EDITED BY
Sara Isernia,
Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy

REVIEWED BY
Ilaria Grazzani,
University of Milano-Bicocca, Italy
Lucia Bigozzi,
University of Florence, Italy

*CORRESPONDENCE
Alessia Cornaggia
✉ cornaggia13@gmail.com

RECEIVED 28 June 2024
ACCEPTED 30 July 2024
PUBLISHED 12 August 2024

CITATION
Cornaggia A, Bianco F, Castelli I and Belacchi C (2024) Assessing metarepresentational abilities in adolescence: an exploratory study on relationships between definitional competence and theory of mind.
Front. Psychol. 15:1456432.
doi: 10.3389/fpsyg.2024.1456432

Assessing metarepresentational abilities in adolescence: an exploratory study on relationships between definitional competence and theory of mind

Alessia Cornaggia[1]*, Federica Bianco[1], Ilaria Castelli[1] and Carmen Belacchi[2]

[1]Department of Human and Social Sciences, University of Bergamo, Bergamo, Italy, [2]Department of Communication Sciences, Humanities and International Studies, University of Urbino Carlo Bo, Urbino, Italy

Introduction: Several developmental changes occur in adolescence, particularly in the metarepresentational domain, which allows and promotes adaptive sociality. We explored the possible relationships between theory of mind (ToM) and definitional competence, both metarepresentational, beyond age and gender effects.

Methods: To reach our goals, we involved 75 adolescents (age range 14–19 years, M = 15.7, and SD = 1.36). ToM was measured through "The Reading the Mind in the Eyes Test" (RMET), and definitional competence was assessed through a new instrument, namely, the "Co.De. Scale". Attention was paid to check whether results were different when considering mental states vs. non-mental states of the scale and emotional words vs. non-emotional words.

Results: *T-tests* showed that older adolescents (third grade of high school) performed better than younger ones (first grade of high school) in both tasks. Only in the male group, there were no school grade differences in the ToM task. Regression analyses showed that RMET performance predicted the score of *non-emotional mental states* definitions and, even if marginally, of *ToM word* definitions. However, RMET was not a predictor of the general performance of the definitional task or *emotion* definitions.

Discussion: Connections with global adolescents' development and possible educational implications are discussed.

KEYWORDS
theory of mind, mindreading, definitional competence scale, adolescence, metarepresentational abilities

1 Introduction

Adolescence is a developmental phase characterized by several changes (Lerner and Steinberg, 2009; Waite-Jones and Rodriguez, 2022), such as major physical transformation, the increase of social relationships, in particular, the relevance of the bond with peers (Carpendale and Lewis, 2004; Zerwas et al., 2004), and also deep cognitive maturation (Byrnes, 2003). In particular, the emotional world in adolescence intertwines with the effects of physical changes that could elicit anxiety, uncertainty, and insecurity; moreover, the experience of these feelings enables self-reflection based on the relational exchanges, especially in the peer group (Palmonari and Crocetti, 2011). This developmental phase is also characterized by complex interactions between interpersonal emotional states, such as guilt, shame, forgiveness, gratitude, self-compassion, and prosocial behavior (Carlo et al., 2023).

Moreover, representational abilities play a crucial role, as identity emerges from the mental representations about oneself and others and also from the meanings attributed to past, present, and possible future experiences (Białecka-Pikul et al., 2020; Bosacki et al., 2020). An important feature of thought in adolescence is the development of metacognitive abilities that consist of "making use of knowledge to remember, reason, make decision and solve problems" (Byrnes, 2003, p. 241), but also the capacity to reflect on one's own and others' knowledge (Moshman, 1998). The adolescents' exercise of thought elaboration allows them to develop personal beliefs, values, and critical abilities that permit them to adapt to their cultural context (Białecka-Pikul et al., 2020).

In this complex framework of changes that involve the acquisition of more abstract levels of representation, in particular, of a metacognitive, metarepresentational, and metalinguistic type, a crucial maturation concerns neural structures and connectivity (Byrnes, 2003; Devine and Lecce, 2021; Laghi and Lonigro, 2022). Although experience could be an input for neural changes, the maturation of frontal lobes (Byrnes, 2003) and the reorganization of the pre-frontal cortex (Blakemore, 2008) seem to specifically contribute to the development of executive functions and metarepresentational abilities, which significantly contribute to the increase of abstract thinking skills during adolescence (Steinberg, 2005).

The type of thought that has as its object the representations of themselves is precisely what characterizes the hypothetical-deductive thought that opens adolescents' cognitive development to the world of possibilities and inferences (Piaget and Inhelder, 1955). Indeed, what distinguishes Piaget's formal operational stage is the ability to operate on abstract ideas and knowledge, due to the independence of thought from current action. Although Piaget's stadial theory does not attribute a primary role to social components in the early stages of development, exposure to others' perspectives is considered an important element in accessing formal operational thinking (Marchetti and Massaro, 2002). In this process of discovering one's own and others' perspectives, Piaget seems to recognize language's function of triggering reflection about beliefs and mental content. Subsequent evolutions of cognitive psychology have revealed how both adolescents and adults do not follow perfect formal logic in their reasoning, but there are systematic errors that could arise from the mental representations and interpretations of premises in a specific socio-cultural context, as well as from the limits of working memory or linguistic pragmatics (Carugati and Selleri, 2011).

The ability to understand that inferences on mental states allow the prediction and possible explanations of behaviors (Premack and Woodruff, 1978; Wimmer and Perner, 1983), namely ToM is a possible crucial factor intertwined with other metarepresentational abilities that affect the main cognitive and affective domains of adolescents' experience (Apperly, 2021; Devine and Lecce, 2021).

1.1 ToM improvement in adolescence

While traditionally research on ToM has focused primarily on the preschool and school-age periods with the aim of overcoming classical false belief tasks (Wellman, 2012; Castelli et al., 2022), across the past two decades, ToM research has turned into a lifespan perspective (Kuhn, 2000; Marchetti et al., 2016; Peterson and Wellman, 2019), showing that ToM ability continues to undergo relevant changes both on the behavioral and on the neural levels (Castelli et al., 2010; Cabinio et al., 2015). Previous studies also explored associations between ToM and social relations (Sebastian, 2015; Lebedeva et al., 2023), between ToM and the socio-emotional domain (Clifford et al., 2021; Mulvey et al., 2022), and between ToM and both neural (Sebastian et al., 2012; Vetter et al., 2014) and cognitive (Altgassen et al., 2014; Wang et al., 2021) domains during adolescence.

A field of research in the literature on ToM in adolescence is concerned with the validity of measurement tools and measures to assess ToM in adolescence (Hayward and Homer, 2017), such as the Animated Triangle Task (Andersen et al., 2022), the Theory of Mind Assessment Scale (Bosco et al., 2014), EmpaToM—Youth (Breil et al., 2021), and the automated ToM measurement through machine learning and deep learning systems (Devine et al., 2023).

If we consider the most traditional measures of ToM, such as the false belief tasks, Valle et al. (2015) highlighted the development of third-order recursive thinking from adolescence to adulthood. Indeed, in participants aged 14, 17, and 20 years, an age effect on their performances at the third level false belief task was found, controlling for general cognitive abilities. Meanwhile, through the Imposing Memory Task (Kinderman et al., 1998), a set of five stories for advanced ToM that involve the recursive thinking ability within complex social situations and one control story detected a correlation between the third level of reasoning and language comprehension ability. The absence of correlations between lower and upper levels revealed the mastery of first- and second-order reasoning, and the great difficulty with the fourth and fifth levels of reasoning (Valle et al., 2015). The development of recursive thinking in late childhood and adolescence was investigated longitudinally by Van Den Bos et al. (2016) on participants aged from 8 to 17 years. The authors observed both cross-sectionally and longitudinally (after 2 years) an effect of age on the recursive thinking ability that seemed to follow a linear development up to the age of 18 years. Moreover, verbal reasoning could play a role in the development of recursive thinking more than vocabulary (Van Den Bos et al., 2016). These results were in line with a previous study by Dumontheil et al. (2010) in a sample of females aged from 7 to 27 years, where researchers highlighted that the perspective-taking ability continued its improvement even in late adolescence when the development of executive functions had already reached the levels of adulthood (Dumontheil et al., 2010). Based on the measures of this study, Symeonidou et al. (2016) used eye-tracking in a sample composed of children, adolescents, and adults, detecting a difference in the adults' ToM performances when compared with younger participants, a result that seems to confirm adolescence as an ongoing developmental phase in reasoning about own and others mental states.

Beyond these results, which mainly concern the cognitive aspects of ToM, the research on ToM development in adolescence also focused on the affective components of mental state reasoning and its social use (Bosacki, 2015). In particular, the longitudinal results by Bialecka-Pikul et al. (2020) in a sample of 13- and 16-year-old Polish adolescents showed an increase in psychological self-descriptions from early to middle adolescence and also a between-group significant difference in advanced ToM. A measure frequently used to assess affective ToM in its socio-perceptual component in the lifespan is the RMET (Baron-Cohen et al., 1997, 2001). Such measures showed significant results in studies with adolescents (Meinhardt-Injac et al., 2020; Gabriel et al., 2021). Meinhardt-Injac et al. (2020) detected a specific effect of age on ToM, which was not related to improvements in other cognitive abilities, such as language and executive functions, in participants aged between 11 and 25 years. More specifically, Gabriel et al. (2021) focused on the development of both affective and cognitive ToM in three phases of adolescence: early (13–14 years), middle (15–16 years), and late (17–18 years) adolescence. The results showed lower performance in both ToM components by early adolescents, whereas no significant changes emerged when comparing between middle and late adolescents. Moreover, in the first period of adolescence, a relationship between affective ToM and verbal abilities, such as fluency, flexibility, and verbal intelligence, was observed, while cognitive ToM was found to be related to language comprehension in all the considered age groups (Gabriel et al., 2021). This study also highlighted better performances in cognitive ToM in female participants compared with male participants. Gender differences in ToM in adolescence are still controversial: evidence in favor of female students has been found in advanced ToM measures (Bosco et al., 2014; Białecka-Pikul et al., 2017, 2020; Gabriel et al., 2021), but they have not always been confirmed (Bosacki et al., 2020). Moreover, possible influences of gender stereotypes, such as a higher social awareness of girls than boys, have to be considered (Bosco et al., 2014; Białecka-Pikul et al., 2021). The study conducted by Białecka-Pikul et al. (2021) also detected an association between advanced ToM and verbal abilities that was particularly evident in female participants. Bosco et al. (2014), using the Theory of Mind Assessment Scale (Th.o.m.a.s.—Bosco et al., 2009), highlighted better performances by female participants in a sample of pre-adolescents and adolescents from 11 to 17 years. In line with the previous illustrated studies, the authors showed an effect of age on ToM development that became more stable in participants older than 15 years (Bosco et al., 2014). Moreover, the longitudinal study conducted by Stępień-Nycz et al. (2021) explored changes in advanced ToM in 13- and 16-year-old participants, observing a female better performance in advanced ToM measured through the Ambiguous Story Task (Bosacki et al., 2015), controlling for language.

As mentioned in previous studies, another important variable that interacts with ToM development is language in its different components and skills (Antonietti et al., 2006; Siegal and Surian, 2011; Pinto et al., 2017; de Villiers, 2021). The studies on first- and second-order ToM development have highlighted a reciprocal relationship between these two areas in the course of development (Belacchi, 2022; Miller, 2022), but in adolescence, the picture is far from clear. The lower number of studies in advanced or mature ToM combined with the increasing complexity of ToM abilities in middle childhood and adolescence does not help to clarify the relationship between ToM and language in adolescents (Milligan et al., 2007; Devine and Lecce, 2021), even if some results have suggested the possible persistence, even at more advanced levels of development, of some kind of relationship between these two abilities (Antonietti et al., 2006; Im-Bolter et al., 2016; Gabriel et al., 2021). In the complex relationship between language and ToM, another relevant variable to consider is mental-state language, which includes those terms that refer to the cognitive, emotional, and volitional spheres (Lecce, 2009). Studies on middle-childhood ToM development detected a significant association between engagement in mental-state conversations with peers or teachers and ToM abilities in students (Ornaghi et al., 2014; Bianco and Lecce, 2016; Bianco et al., 2019, 2021; Lombardi et al., 2022; Bianco and Castelli, 2023). Indeed, mental-state language allows both the observation of the first expressions of ToM and the understanding of the interactional development of different metarepresentational abilities (Meins et al., 2006; Lecce and Pagnin, 2007). Im-Bolter et al. (2016), in a study that involved participants aged between 7 and 12 years old, detected two different models of interactions between higher-order ToM and other cognitive domains, respectively, for middle childhood and early adolescence. Regarding this second life period, the authors highlighted a less complex model characterized in particular by a lower involvement of mental attentional capacity measured in both its verbal and visuospatial components. This is probably due to the greater experience with ToM reasoning and the development of higher cognitive skills. Moreover, they found a decreased relevance of syntactic language abilities (a closed system), in favor of a persistent role played by the semantic component (an open system), considered a language competence with a higher developmental potential in adolescence and adulthood. Some studies also investigated the possible connections between language and ToM, considering even the social components; for instance, Brodsky et al. (2023) detected a role of language in adolescents' performance in interpreting unambiguous social scenarios. Lavoie and Talwar (2022) highlighted the role of ToM in determining the tendency of adolescents to maintain transparency and sharing of information with both parents and friends. Widening the perspective, Pluck et al. (2021) investigated the role of adolescents' SES on language ability, related to ToM and executive function, showing a stronger association with language than with ToM or executive function. The association between ToM and executive function also became non-significant when language was added as a control variable.

In line with these suggestions, language and ToM could be considered important metarepresentational functions that deserve to be further investigated in adolescence, as they evolve in the interaction with the set of complex changes and acquisitions that characterize this phase of life.

1.2 Definitional competence in adolescence

Language plays a key role in adolescence: expressing one's own point of view, discussing, arguing, and exercising critical thinking are crucial skills (Tolchinsky and Berman, 2023). Therefore, metalinguistic and metarepresentational functions, specifically expressed through definitional activity, become fundamental, especially because they allow using language to reflect on language itself in order to share, in a decontextualized way, the semantic representations with anyone who knows our language (Belacchi, 2022). In particular, producing a lexicographic definition, such as a sentence that expresses the meaning (*definiens*) of a given word (*definiendum*), requires combining content and form components (Benelli et al., 2006), but above all, a de-contextualizational perspective that takes into account the interpersonal culturally shared meaning of words (Benelli et al., 2006; Belacchi and Benelli, 2017, 2021). Conventionally, a lexicographic definition, in order to pursue the necessary communicative effectiveness, must assume the Aristotelian format: "An X is a Y that

Z," where "X" represents the given object or concept, "Y" represents the *genus proximum* (the superordinate category), and Z represents the *differentia specifica*, i.e., information that allows the specific object or concept to be identified (Benelli et al., 2006). In particular, a lexicographic definition must have five requisites or rules: a correct and syntactically autonomous linguistic structure (*correctness and morpho-syntactic autonomy rule*), a semantic correspondence between the stimulus item and the sentence that explains its meaning (*semantic equivalence rule*), expressed by verbs such as, "means" "refers to," "is" (*copula rule*), an articulated sentence (*periphrasis* or *phrasal extension rule*) and the use of different words from the *definiendum* (*no-tautology rule*) (Benelli et al., 2006; Belacchi and Benelli, 2017, 2021). In order to reach such a high level of complex skills, both of form and content, as those required in a definitional task, a process of development in a bidirectional influence with formal instruction is required (Benelli et al., 2006; Artuso et al., 2021). The developmental trend of definitional competence and its specific components can be assessed through the Competence Definitional Scale, Co.De. Scale, (Benelli et al., 2006; Belacchi and Benelli, 2017), whose last version (Belacchi and Benelli, 2021) is structured in seven progressive levels that reflect different degrees of definitional ability. Refer to Table 1 for more details on the levels of the Co.De. Scale. The recent normative study of the development of definitional competence in the Italian population through the Co.De. Scale (Belacchi and Benelli, 2021) showed a significant age improvement (from preschoolers to adults) without significant gender differences. Of note, Dourou et al. (2020), assessing the ability to define words from preschoolers to adults, found a vantage in the female participants in all age groups.

It has been well-documented that definitional competence undergoes significant changes across the lifespan and education levels. The studies by Belacchi and Benelli (2007, 2017, 2021), starting from the first proposals by Litowitz (1977), Watson (1985, 1995), Johnson and Anglin (1995), and Nippold (1995), have detected a development in the quality of definitions from the preschool age to adulthood. Around the age of 7 years, children start to systematically use superordinate categorical terms ("ISA structure") and not only the so-called "HAS structure" descriptive definitions that characterize the preschool years (Benelli et al., 2006). Around the age of 10–12 years, the quality of definitions becomes similar to the one at adult levels (Benelli et al., 2006). Interestingly, the 11-year-old participants performed better than adults with low educational levels, suggesting the possible role of daily practice with definitions and school learning in determining the quality of definitions (Snow et al., 1989; Snow, 1990). A relevant contribution to the understanding of the developmental process of definitional competence was provided by Belacchi and Benelli (2017) with a study that involved participants aged from 5 to 20 years, divided into five age groups. The results showed not only an increase in lexical knowledge, but also a growing appropriateness, completeness, and formal correctness in comparing all the considered age groups. Recently, a study conducted on the elderly population showed a decline in definitional competence associated with aging (Bianco et al., 2022) and also in the taxonomic organization of representations, a fundamental ability for definitional competence (Belacchi and Artuso, 2018). The study by Bianco et al. (2022) also showed how ToM and definitional competence are related to aging. In particular, ToM was detected as a predictor of definitional competence; moreover, that work showed that definitional competence seems to decline earlier than ToM. In this study, ToM was measured using the Reading the Mind in the Eyes Test (Baron-Cohen et al., 1997, 2001) and the definitional competence through the Co.De. Scale (Belacchi and Benelli, 2021).

Important implications of metarepresentational skills have also been observed in relational and social development, particularly in dealing with the complex phenomenon of bullying (Belacchi and Benelli, 2020). In participants aged between 8 and 10 years, lower levels of definitional competence, particularly related to emotion terms, were associated with hostile roles and predicted aggressive behavior (Belacchi and Benelli, 2020). Moreover, in this study, a positive correlation between two different metarepresentational skills (definitional competence and the cognitive component of empathy) was observed. The common meta-representational nature of cognitive, affective, and relational skills, as well as their influence on social dynamics, are issues that may provide interesting insight into development in adolescence, when relationships, especially with peers, are fundamental for wellbeing (Bagwell and Schmidt, 2013).

In this view, and to the best of our knowledge, the present study focused on definitional competence—a peculiar expression of metalinguistic and metarepresentational skills—in adolescence and on its possible connections with ToM.

As suggested by this theoretical framework and previous studies in the literature, the development of metarepresentational skills (i.e., ToM and definitional competence) could play a crucial role in the significant changes that occur in cognitive, emotional, and social domains during adolescence. Indeed, the ability to understand, reflect, and share mental representations about the connections between internal states and reality was associated with

TABLE 1 Definitional levels, prototypical answers, and scores for the definition of the word "Donkey."

Levels	Kinds of answers	Score
0. Non-definition	Non-verbal answers	0
I. Pre-definition	One-word answers, mostly associations (e.g., *donkey- > ears*)	1
II. Quasi-definition	The initial formulation of sentences, without autonomous forms (e.g., *donkey- > with the long ears; when it brays*)	2
III. Narrative/descriptive definition	Formally correct and autonomous sentences, with narrative/descriptive content (e.g., *donkey brays; donkey is mild*)	3
IV. Categorical definition	Formally correct and autonomous sentences in simply categorical/ synonymic form (e.g., *The donkey is an animal*)	4
V. Partial aristotelian definition	Formal correctness without semantic equivalence (e.g., *The moon is a planet in the solar system*)	5
VI. Aristotelian, metalinguistic definition	Formal and semantic correctness and equivalence (e.g., *A donkey is an animal that brays*)	6

self-understanding processes (Białecka-Pikul et al., 2020), such as the quality of friendship, relationships (Fink, 2021), and social behavior (Bosacki, 2021) in adolescence. Moreover, definitional competence has some connections with social outcomes in children aged between 8 and 10 years (Belacchi and Benelli, 2020), which encourages exploring its connection with other developmental domains in adolescence.

1.3 Research questions and hypothesis

The present study aims to deepen our understanding of the development of metarepresentational abilities in adolescence, in particular of a component of the ToM construct, i.e., the ability to infer mental states from eye images, and the definitional competence (regarding both ToM and non-ToM words) as an expression of metalinguistic and metarepresentational abilities.

The first specific aim is to investigate possible differences in the development of ToM and of definitional competence between early and middle adolescence, due to the absence of studies in literature focused on this development phase, to the best of our knowledge. The plurality of changes that occur in adolescence, as shown in previous literature, could be more stabilized in middle adolescence: therefore, we expect better performances from the older group on both tasks measuring the two abilities.

Second, following the evidence concerning gender differences in development during adolescence, we explore possible gender specificities. Despite the absence of corroborating evidence in the existing literature, some suggestions support the hypothesis that female subjects may exhibit superior performance compared to their male counterparts (Bosco et al., 2014; Longobardi et al., 2016; Gabriel et al., 2021).

Finally, we investigate the relationship between the two metarepresentational abilities involved in the study, and, in line with Bianco et al. (2022), we hypothesize that ToM could be a predictor of definitional competence, especially in defining ToM words.

2 Materials and methods

2.1 Participants and research context

The research involved 75 adolescents (33 males and 42 females) aged from 14 to 19 years ($M = 15.7$ $SD = 1.36$). The research was conducted in the center of Italy, involving three different public high schools: lyceum ($N = 31$), the technical institute ($N = 19$), and the professional institute ($N = 25$). Inclusion criteria were fluency in the Italian language, the absence of neuro-developmental or psychiatric disorders, and attendance in the first and third grades of high school. A total of 36 participants attended the first grade and 39 attended the third grade. The classes were selected on the proposal of school leaders, who assessed the willingness of teaching staff to collaboratively engage. The informed written parental consent and participant's consent were obtained. Data were collected in the first part of the school year. All requirements of ethical guidelines provided by the Declaration of Helsinki (World Medical Association, 2013), the American Psychological Association (APA, 2017), and the Italian Psychological Association (AIP, 2022) were amended.

2.2 Measures

2.2.1 Theory of mind

The Reading the Mind in the Eyes Test (RMET) (Baron-Cohen et al., 1997, 2001) in its Italian version (Serafin and Surian, 2004) was chosen to assess the ability to infer complex emotional and epistemic mental states from images of the eye area of human faces. A total of 18 stimuli ($\alpha = 0.341$)[1] were presented to participants and subjects were required to select which mental state was represented in the eyes image. For each image, four options, in the form of a single word, were displayed. One point was given for each correct answer. For instance, when looking at an eyes picture, participants have to choose the correct answer: *giocoso* (playful) excluding the wrong labels *confortante* (comforting) *irritato* (irritated) *annoiato* (bored). The total score ranged from 0 to 18.

2.2.2 Definitional competence

The Co.De. Scale (Belacchi and Benelli, 2021) is used to assess seven progressive levels in the ability to define 32 different target words ($\alpha = 0.884$). Participants were required to provide the meaning of eight nouns (caring, spying, rivalry, donkey, clown, orange, skill, and umbrella), eight adjectives (innocent, thin, round, risky, polite, blond, smooth, and contagious), eight verbs (to think, to tolerate, to force, to frustrate, to beat, to burn, to join, and to emigrate), equally distributed between abstract and concrete terms, and eight terms referring to emotions (pride, sadness, anger, shame, envy, guilt, joy, and fear). Each written answer is attributed scores from 0 (non-definition level) to 6 (Aristotelian, metalinguistic definition) along the following scale (an example of response and related score is illustrated in Table 1). After assigning a score (ranging from 0 to 6), to each answer, we calculated the sum of different scores: a general measure or A "*Total*" score that comprised all items, and some more specific scores that were created by distinguishing "*Non-ToM words*" (donkey, to beat, to burn, clown, thin, round, risky, to join, orange, polite, to emigrate, blond, skill, smooth, umbrella, and contagious) from "*ToM words*" (pride, sadness, anger, shame, envy, guilt, joy, fear, caring, innocent, to think, spying, to tolerate, to force, rivalry, and to frustrate), as made in a previous study in the same domain (Bianco et al., 2022). In order to better understand the connections between definitional competence and the multicomponential nature of ToM, the other two scores were calculated, by separating words describing "*Emotions*" (pride, sadness, anger, shame, envy, guilt, joy, and fear) from words referring to "*Non-emotional mental states*" (caring, innocent, to think, spying, to tolerate, to force, rivalry, and to frustrate). In doing that, we have followed the mental-state language categorization proposed by Lecce and Pagnin (2007).

2.3 Procedure

Data collection was performed at school. Tasks were collectively administered in classrooms under the presence of the researcher in

1 The low index for RMET, a widely used measure in ToM domain, is in line with previous studies (Olderbak et al., 2015; Hayward and Homer, 2017).

one session lasting approximately 1 h. To guarantee that each individual was able to work autonomously, the desks were positioned at a distance from one another. Answers were provided in written format. Participants were reassured about the absence of any form of evaluation. The RMET was the first task presented, followed by the definitional competence task.

2.4 Analysis

The collected data have been analyzed using Jamovi version 1.6.23 statistical software (The Jamovi Project, 2022). Independent-sample t-tests were used to examine the possible effects of school grade and gender on both definitional competence and ToM. Pearson's correlations and partial correlations (inserting age as a control variable) were used to explore possible relations between the ToM component measured by RMET and each of the Co.De. Scale scores. To better understand the possible associations between ToM and definitional competence, hierarchical regressions were performed, entering at Step 1 age and, at Step 2, the hypothesized predictor.

3 Results

Table 2 presents the descriptive statistics of the considered measures. After we consider age and gender differences in definitional competence and ToM tasks, the last section analyzes the relationships between the two abilities.

3.1 Definitional competence and ToM: age and gender differences

As shown in Table 3, all scores showed better performances by the third-grade students compared with the first-grade students. However, results seemed to suggest an association between age and gender because, considering separately male and female groups (Table 4), in the female group, third-grade students performed better than first-grade students in both metarepresentational measures included in the present study ($p = 0.019$, for RMET and $p = 0.001$, for the Co.De. Scale). In the male group, the only significant difference between first- and third-grade students was the better performance of older male participants in *non-emotional mental states* score, $p = 0.041$.

3.2 Relation between definitional competence (the Co.De. Scale) and the ability to read mental state from gaze (RMET)

Table 5 illustrates the correlations between RMET and all the scores on the Co.De. Scale. As we can observe, the expected associations between the ToM score and different definition scores, even if all positive, were not significant. Only correlations between ToM and some specific Co.De. scores were statistically significant. In particular, between RMET and *ToM Words* in the Co.De. Scale ($r = 0.271$, $p = 0.019$), and RMET and *Non-emotional mental states* ($r = 0.277$, $p = 0.016$). This last positive correlation remains significant, $r = 0.234$, $p = 0.045$, also adding age as a control variable.

In order to better understand the relationship between the two metarepresentational abilities, we performed a series of regression analyses. Table 6 shows the regression analysis on the Co.De. Scale scores. We entered age at Step 1, adding RMET at Step 2. Regarding the *Total* score of the Co.De. Scale, Step 1 was significant, $F(1, 73) = 6.51$, $p = 0.013$, while Step 2 did not increase the variance explained, $\Delta F(1, 72) = 1.62$, and $p = 0.21$. When we investigated the impact of ToM on definitional competence, separating *non-ToM* and *ToM words*, the pattern of results was different. In *ToM words* score, as seen for the *total* score, Step 1 (i.e., age) was significant, predicting *ToM words*, $F(1, 73) = 7.78$, $p = 0.007$, $\beta = 3.32$, $t = 2.79$, and $p = 0.007$. Step 2 led to a marginally significant increase in the variance explained, $\Delta F(1, 72) = 3.90$, $p = 0.052$. Differently, age and RMET were not predictors of *non-ToM words*. In order to deepen the significant results on *ToM words*, it is interesting to observe the pattern of results of the sub-scores. For what concerns, *non-emotional ToM words* both Step 1, $F(1, 73) = 7.93$, and $p = 0.006$, and Step 2, $F(2, 72) = 6.22$, and $p = 0.003$, were significant, with an increase in the significant variance explained by RMET, $\Delta F(1, 72) = 4.17$, and $p = 0.045$. On the other hand, for *emotions*, only Step 1 was significant [$F(1, 73) = 4.61$, $p = 0.035$], and Step 2 did not increase the variance explained, $\Delta F(1, 72) = 2.14$, $p = 0.15$.

We also investigated the opposite direction, i.e., the Co.De. Scale scores as predictors of RMET. Specifically in Step 1, we entered age, and in Step 2, the scores of the Co.De. Scale in each regression analysis were entered. In this case, Step 1 was never significant, $p_s \geq 0.103$. Interestingly, *ToM words* Co.De. Scale (marginally), $\beta = 0.23$, $t = 1.98$, and $p = 0.052$, and *non-emotional ToM words* Co.De. Scale, $\beta = 0.24$, $t = 2.04$, and $p = 0.045$, both predicted the RMET score, with the following increases in variance: $\Delta R^2 = 0.05$. Other relations were all non-significant.

TABLE 2 Descriptive statistics of RMET and the Co.De. Scale measures (total, non-ToM words, ToM words, emotions, and non-emotional mental states).

	RMET	Co.De. total	Non-ToM words	ToM words	Emotions	Non- emotional mental states
M	11.3	117	59.2	58.1	31.6	28.9
SD	2.14	20.2	8.26	13	8.35	7.86
Min	7	75	42	33	13	7
Max	15	160	76	84	45	44
Sk	−0.04	0.01	0.05	0.08	−0.478	−0.44
Ku	−0.35	−0.30	−0.54	−0.49	−0.832	0.34

TABLE 3 Independent sample *t*-test on RMET and the Co.De. Scale scores measured in two school grades.

	Grade 1 (*N* = 36)		Grade 3 (*N* = 39)		Student *t*	*df*
	M	*SD*	*M*	*SD*		
RMET	10.81	2.55	11.95	2.55	−2.13*	73
Co.De. total	112.81	20.57	129.23	20.03	−3.50***	73
Non-ToM words	58.53	7.85	62.87	8.96	−2.23*	73
ToM words	54.28	14.11	66.36	12.39	−3.95***	73
Emotions	28.64	8.54	34.36	7.25	−3.14**	73
Non-emotional mental states	25.64	7.40	32	7.06	−3.81***	73

Welch t, equal variances not assumed. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TABLE 4 Independent sample *t*-test on RMET and the Co.De. Scale scores in two gender groups.

	Grade 1 (*N* = 36)			Grade 3 (*N* = 39)			Student *t*	*df*
	N	*M*	*SD*	*N*	*M*	*SD*		
RMET								
Girls	16	10.3	2.41	26	11.5	2.82	−2.44 *	40
Boys	20	11.2	1.64	13	12.2	2.43	- 0.30	31
Co.De. total								
Girls	16	110.88	22.43	26	132.96	18.69	−3.45***	40
Boys	20	114.35	19.41	13	121.77	21.27	−1.03	31
Non-ToM words								
Girls	16	58.13	7.51	26	64.42	8.92	−2.35*	40
Boys	20	58.85	8.29	13	59.77	8.53	−0.308	31
ToM words								
Girls	16	52.75	16.60	26	68.54	11.20	−3.69***	40
Boys	20	55.50	12.07	13	62	13.92	−1.42	31
Emotions								
Girls	16	27.2	9.47	26	36.2	5.88	−3.79***	40
Boys	20	29.8	7.77	13	30.8	8.56	−0.337	31
Non-emotional mental states								
Girls	16	25.6	8.52	26	32.4	6.57	−2.92**	40
Boys	20	25.7	6.61	13	31.2	8.20	−2.14*	31

Welch t, equal variances not assumed. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

To sum up, age was a predictor of *Total* Co.De. Scale score and sub-scores: *ToM words, non-emotional ToM words,* and *emotions*. RMET was a predictor of *non-emotional ToM words* and there was a marginally significant result for the *ToM words* sub-score of the Co.De. Scale. Considering the opposite direction, *ToM words* and *non-emotional ToM words* sub-scores of the Co.De. Scale were, respectively, marginally significant and significant predictors of RMET.

4 Discussion

This study, though preliminary, aims to explore the relations between ToM and definitional competence in adolescence, integrating the developmental processes of metarepresentational abilities into the broader complexity of changes that lead adolescents to build their personal and social identities. There are no previous studies specifically investigating definitional competence in adolescence through the Co.De. Scale or its relationship with ToM in this period of life.

In line with what was hypothesized, the results showed better performances of older groups of participants both in RMET and in all scores of the Co.De. Scale. The significance of age as a predictor for multiple Co.De. Scale scores could depend on multiple factors. One of these could be the role played by formal education in the improvement of students' reasoning and language skills (Benelli et al., 2006; Dourou et al., 2020; Artuso et al., 2021). In particular, third-grade students could be already more familiar with the different metarepresentational abilities involved in high school grade education, which for first-grade students could be an ongoing challenge to deal with (Byrnes, 2003). Another factor could be represented by brain structural maturation, with increases in connectivity in the prefrontal, temporal, and temporoparietal areas that characterize the adolescence

TABLE 5 Bivariate and partial correlations (weighted for age) among RMET and the Co.De. Scale measures.

	RMET	Co.De. total	Non-ToM words	ToM words	Emotions	Non-emotional mental states
RMET	1	0.194	0.035	0.271*	0.208	0.277*
Co.De. total	0.148	1	0.900***	0.965***	0.854***	0.872***
Non-ToM words	−0.003	0.897***	1	0.755***	0.641***	0.710***
ToM words	0.227+	0.962***	0.744***	1	0.900***	0.886***
Emotions	0.170	0.844***	0.623***	0.894***	1	0.596***
Non-emotional mental states	0.234*	0.860***	0.696***	0.874***	0.565***	1

In the lower part of the matrix, we have the partial correlations (weighted for age). $^{+}<0.1$, $*p<0.05$, $**p<0.01$, $***p<0.001$.

TABLE 6 Hierarchical regressions predicting definitional competence.

		R^2	*B*	*SE*	*β*	*t*	*p*
Co.De. total							
Step 1	Age	0.08	4.60	1.80	0.29	2.55	0.013
Step 2	Age	0.10	4.16	1.83	0.26	2.27	0.026
	RMET		1.33	1.04	0.15	1.27	0.207
Non-ToM words							
Step 1	Age	0.04	1.28	0.73	0.20	1.75	0.085
Step 2	Age	0.04	1.28	0.75	0.20	1.71	0.092
	RMET		−0.01	0.43	−0.00	−0.03	0.979
ToM words							
Step 1	Age	0.10	3.32	1.19	0.31	2.79	0.007
Step 2	Age	0.14	2.87	1.19	0.27	2.42	0.018
	RMET		1.34	0.68	0.22	1.98	0.052
Emotions							
Step 1	Age	0.06	1.50	0.70	0.24	2.15	0.035
Step 2	Age	0.09	1.31	0.71	0.21	1.85	0.069
	RMET		0.59	0.40	0.17	1.46	0.148
Non-emotional mental states							
Step 1	Age	0.10	1.82	0.65	0.31	2.82	0.006
Step 2	Age	0.15	1.57	0.64	0.27	2.44	0.017
	RMET		0.75	0.37	0.23	2.04	0.045

phase (Devine and Lecce, 2021; Gabriel et al., 2021; Laghi and Longiro, 2022). Moreover, as suggested by other evidence that detected differences between early and middle adolescence, the higher tendency of early adolescents to prioritize self-perception over the effort to understand others (Bosco et al., 2014) could have affected, in the younger group, both the ability to infer mental states from the eyes of another person and the accuracy in providing clear information in defining terms. Our findings about ToM are in line with those of Gabriel et al. (2021), who observed a significant increase in both cognitive and affective ToM between 13- and 16-year-olds, suggesting a higher cognitive demand for younger participants, especially in terms of attention. In addition, Bialecka-Pikul et al. (2020), exploring the interactions between advanced ToM and self-construction, highlighted the relation between ToM and the process of self-construction, finding significant improvement in both developmental dimensions in adolescents from 13 to 16 years. Of note, we observed a particularly high value of SD for the Co.De. Scale. This indicates the presence of high variability of performance in our sample for definitional competence. Even if this result is preliminary and should be confirmed in future research, a possible explanation could be related to the relevant changes in cognitive skills that happen during adolescence (Carugati and Selleri, 2011).

The second specific aim of this study was to explore possible gender differences in our sample, given that previous results were rather mixed in their conclusions on this issue (Bosco et al., 2014; Białecka-Pikul et al., 2017; Dourou et al., 2020; Belacchi and Benelli, 2021; Stępień-Nycz et al., 2021; Caputi and Bosacki, 2023). Our results partially confirm the better performances of female participants compared with male participants because, in the ToM task, we observed the above-mentioned age group's significant difference only in the female group. In agreement with what was hypothesized in other studies (Bosco et al., 2014; Gabriel et al., 2021), this difference could be explained by a different development of neuropsychological processes and structures, but also by socio-cultural factors and competences. Even in a study conducted by Longobardi et al. (2016) on mental state language in primary schools, a better performance of female participants in the use of terms related to mental states was detected, although male participants obtained higher scores than female participants in receptive language. In order to better understand these gender differences that are not consistent in the literature, further studies are needed that could confirm these findings and more fully explore the possible variables involved.

The third and more innovative purpose of this study focuses on the possible associations between the two metarepresentational abilities: ToM and definitional competence. Consistent with a similar study conducted by Bianco et al. (2022) in aging, RMET performances also predicted the score of *ToM words* definitions in adolescence. However, in the present study, RMET was neither a predictor of general performance in the definitional task nor on *emotion* definitions. The only definitional scores that have been significantly predicted by the RMET score were the *non-emotional mental states* score and, even if marginally, the *ToM words* score. Reciprocally, none of the definitional scores was a predictor of RMET performance. These results confirm the presence of some subtle, but not negligible, associations between the implicit metarepresentational ability to infer mental states and to produce lexicographic definitions also in adolescence, stressing the link between ToM and language competence, often remarked in literature (de Villiers, 2021). In particular, Gabriel et al. (2021) highlighted the role of verbal ability in facilitating the process of affective ToM reasoning knowing the meaning of emotion words. Similarly, Im-Bolter et al. (2016) highlighted in adolescence the possible interaction between ToM and semantic components of language. As suggested in the previous studies with older people (Bianco et al., 2022), the specific association between RMET and the Co.De. Scale could be in the direction of moving from the implicit nature of the ToM task to the explicit nature of metarepresentational processes necessary to produce a high-level definition. The unexpected lack of significant associations between RMET and the definitions of emotions could be explained by the emotional turbulence that characterizes the adolescent phase, which can make conscious access to emotions, and the consequent conventional verbalization, which is particularly difficult (Marchetti and Cavalli, 2013; Gatta et al., 2014; Muzi, 2020).

This study presents some limitations related to the need to segment the complexity of developmental processes while reading the results obtained within a systemic framework (Bronfenbrenner, 1979) that takes into account the mutual interactions between the multiple domains and sources of development. The first limitation of the study is ToM measurement, which, especially for advanced forms of mental state reasoning, would require the employment of a battery composed of multiple and more articulated tools that would allow for consideration of different components of the ToM construct (Beaudoin et al., 2020), such as the cognitive and affective ToM (Gabriel et al., 2021). Similarly, language should be examined not only in its metalinguistic aspect as a definitional competence but also through the employment of multiple instruments that investigate different linguistic competences. The third important limitation is the sample size; thus, future studies will need the recruitment of a larger sample to derive more accurate conclusions. Finally, a limitation regards the lack of information about participant social and affective life or their specific abilities in other cognitive and emotive domains. Future studies could broaden the perspective in these directions, with the aim of placing the development of metarepresentational abilities within a more comprehensive framework that investigates reciprocal associations of different domains. It could be useful in the future to apply a longitudinal approach to the issues investigated here.

5 Conclusion

In summary, this study deals with the complexity of development in adolescents through the investigation of implicit and explicit levels of metarepresentational abilities. The major innovative element was the exploration of the use of the Co.De. Scale in this specific age period, which added new knowledge to the relationship between metalinguistic skills and ToM. Definitional competence could already represent a construct that allows for an increase in the understanding of the interaction between language and the ability to infer mental states. For this reason, it could be interesting in future studies to include multiple measurements of ToM to explore the possible different connections of definitional competence and different components of ToM. Because of the relevance of social and relational dimensions in adolescence, it could be worth deepening the role of these two metarepresentational abilities in interaction with other key factors of development, such as the quality of social interactions, emotion recognition, and social adjustment. For instance, it may be beneficial to consider ways to support the development of metarepresentational skills as a potential way to assist adolescents in navigating the emotional challenges that accompany the many changes they experience during this period of life. Increasing knowledge about specific metarepresentational abilities and the relationship between them could also set the basis for the implementation of training studies with multiple possible educational implications, as suggested also by Belacchi and Benelli (2020) for what concerns the connections of definitional competence and empathy with bullying roles in primary school children. Adolescence could be a sensitive period for prevention and people's wellbeing, but programs should be based on a detailed understanding of individual differences and should include the promotion of multiple skills (Caputi and Bosacki, 2023). For example, as observed by studies concerning the "Promoting mental health at schools" (PROMEHS) program (Cefai et al., 2022; Colomeischi et al., 2022; Martinsone et al., 2022), processes of social and emotional learning could play a role for adolescents' wellbeing/prosocial behavior, and these competences are also likely to increase resilience in critical situations such as those that occurred during the pandemic. Moreover, the linguistic and metalinguistic abilities that enable an individual to represent and socially share their internal states could be relevant factors, especially for the development of social and emotional learning (Cavioni et al., 2017). We hope our study can help in exploring these directions.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The requirement of ethical approval was waived by IRB Committee University of Urbino for the studies involving humans because Article 8 of the IRB (released with Rectoral Administrative Order n. 29/2019 of the 29th of January 2019, operating from the 1st of February 2019) about Presentation and supervision of research projects states: "1. Research projects that involve experimental collection of human behaviors (e.g., response accuracy, response latencies) and/or personal data, and/or bio-material from human beings, developed and activated within the University of Urbino, do not need to be necessarily submitted to the IRB Committee evaluation." The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

AC: Formal analysis, Writing – original draft, Writing – review & editing. FB: Conceptualization, Formal analysis, Methodology, Supervision, Writing – review & editing. IC: Conceptualization, Methodology, Supervision, Writing – review & editing. CB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We are grateful to all participants for their collaboration and to Irene Roselli for the help in the investigation and data curation.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

AIP (2022). Codice etico per la ricerca in psicologia. Available at: https://aipass.org/commissione-etica (Accessed August 5, 2024).

Altgassen, M., Vetter, N. C., Phillips, L. H., Akgün, C., and Kliegel, M. (2014). Theory of mind and switching predict prospective memory performance in adolescents. *J. Exp. Child Psychol.* 127, 163–175. doi: 10.1016/j.jecp.2014.03.009

Andersen, N. K., Rimvall, M. K., Jeppesen, P., Bentz, M., Jepsen, J. R. M., Clemmensen, L., et al. (2022). A psychometric investigation of the multiple-choice version of animated triangles task to measure theory of mind in adolescence. *PLoS One* 17:e0264319. doi: 10.1371/journal.pone.0264319

Antonietti, A., Liverta Sempio, O., and Marchetti, A. (2006). Theory of mind and language in different developmental contexts. New York: Springer.

APA (2017). Ethical principles of psychologists and code of conduct ethical principles of psychologists and code of conduct. *Ethic. Princ. Psychol. Code Cond.* 57. Available at: https://www.apa.org/ethics/code

Apperly, I. A. (2021). "Cognitive basis of mindreading in middle childhood and adolescence" in *Theory of mind in middle childhood and adolescence: Integrating multiple perspectives*. (eds.) R. T. Devine and S. Lecce, (Abingdon: Routledge), 37–54.

Artuso, C., Palladino, P., Valentini, P., and Belacchi, C. (2021). Definitional skills as a bridge towards school achievement. *Sustain. For.* 14:286. doi: 10.3390/su14010286

Bagwell, C. L., and Schmidt, M. E. (2013). Friendships in childhood and adolescence. New York: Guilford Press.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The "Reading the mind in the eyes" test revised version: a study with Normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1111/1469-7610.00715

Beaudoin, C., Leblanc, É., Gagner, C., and Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Front. Psychol.* 10:2905. doi: 10.3389/fpsyg.2019.02905

Belacchi, C. (2022). "Teoria della Mente e linguaggio Come condividere emozioni e pensieri con le parole" in La competenza sociale nei bambini. Valutare e promuovere la Teoria della Mente in età prescolare e scolare. eds. I. Castelli, D. Massaro and A. Marchetti (Bergamo: Edizioni Junior), 101–134.

Belacchi, C., and Artuso, C. (2018). How taxonomic and thematic associations in semantic memory modulate recall in young through old-old adults. *Psychol. Aging* 33, 1060–1069. doi: 10.1037/pag0000297

Belacchi, C., and Benelli, B. (2007). *Il significato delle parole. La competenza definitoria nello sviluppo tipico e atipico*. Bologna: Il Mulino.

Belacchi, C., and Benelli, B. (2017). A mosaic is not a place where you go to see pictures: on the nature of definitions and the development of definitional skills. *Riv. Psicolinguist. Appl.* 17, 71–91. doi: 10.19272/201707701004

Belacchi, C., and Benelli, B. (2020). Meta-representational skills in bullying roles: the influence of definitional competence and empathy. *Front. Psychol.* 11:592959. doi: 10.3389/fpsyg.2020.592959

Belacchi, C., and Benelli, B. (2021). Valutare la competenza definitoria. La Scala Co. De. in ambito clinico e nello sviluppo tipico. Milan: Franco Angeli.

Benelli, B., Belacchi, C., Gini, G., and Lucangeli, D. (2006). 'To define means to say what you know about things': the development of definitional skills as metalinguistic acquisition. *J. Child Lang.* 33, 71–97. doi: 10.1017/S0305000905007312

Białecka-Pikul, M., Kołodziejczyk, A., and Bosacki, S. (2017). Advanced theory of mind in adolescence: do age, gender and friendship style play a role? *J. Adolesc.* 56, 145–156. doi: 10.1016/j.adolescence.2017.02.009

Białecka-Pikul, M., Stępień-Nycz, M., Szpak, M., Grygiel, P., Bosacki, S., Devine, R. T., et al. (2021). Theory of mind and peer attachment in adolescence. *J. Res. Adolesc.* 31, 1202–1217. doi: 10.1111/jora.12630

Białecka-Pikul, M., Szpak, M., Zubek, J., Bosacki, S., and Kołodziejczyk, A. (2020). The psychological self and advanced theory of mind in adolescence. *Self Identity* 19, 85–104. doi: 10.1080/15298868.2018.1538900

Bianco, F., and Castelli, I. (2023). The promotion of mature theory of mind skills in educational settings: a mini-review. *Front. Psychol.* 14:1197328. doi: 10.3389/fpsyg.2023.1197328

Bianco, F., Castelli, I., and Belacchi, C. (2022). Changes of Meta-representational skills in ageing: first empirical evidence on the relation between metalinguistic competence and attributions of mental states. *J. Lang. Educ.* 8, 40–51. doi: 10.17323/jle.2022.13868

Bianco, F., and Lecce, S. (2016). Translating child development research into practice: can teachers foster children's theory of mind in primary school? *Br. J. Educ. Psychol.* 86, 592–605. doi: 10.1111/bjep.12125

Bianco, F., Lombardi, E., Lecce, S., Marchetti, A., Massaro, D., Valle, A., et al. (2021). Supporting Children's second-order recursive thinking and advanced ToM abilities: a training study. *J. Cogn. Dev.* 22, 561–584. doi: 10.1080/15248372.2021.1901712

Bianco, F., Lombardi, E., Massaro, D., Castelli, I., Valle, A., Marchetti, A., et al. (2019). Enhancing advanced theory of mind skills in primary school: a training study with 7- to 8-year-olds. *Infant Child Dev.* 28:e2155. doi: 10.1002/icd.2155

Blakemore, S.-J. (2008). The social brain in adolescence. *Nat. Rev. Neurosci.* 9, 267–277. doi: 10.1038/nrn2353

Bosacki, S. L. (2015). Children's theory of mind, self-perceptions, and peer relations: a longitudinal study. *Infant Child Dev.* 24, 175–188. doi: 10.1002/icd.1878

Bosacki, S. (2021). "Theory of mind and peer relationships in middle childhood and adolescence" in *Theory of mind in middle childhood and adolescence* (eds.) R. T. Devine and S. Lecce, (Abingdon: Routledge), 142–168.

Bosacki, S., Bialecka-Pikul, M., and Szpak, M. (2015). The adolescent mind in school: theory of mind and self-concept in Canadian and polish youth. *Int. J. Adolesc. Youth* 20, 457–469. doi: 10.1080/02673843.2013.804423

Bosacki, S., Moreira, F. P., Sitnik, V., Andrews, K., and Talwar, V. (2020). Theory of mind, self-knowledge, and perceptions of loneliness in emerging adolescents. *J. Genet. Psychol.* 181, 14–31. doi: 10.1080/00221325.2019.1687418

Bosco, F. M., Colle, L., Fazio, S.De, Bono, A., Ruberti, S., and Tirassa, M. (2009). Th.O.M.A.S: an exploratory assessment of theory of mind in schizophrenic subjects. *Conscious. Cogn.* 18, 306–319. doi: 10.1016/j.concog.2008.06.006

Bosco, F. M., Gabbatore, I., and Tirassa, M. (2014). A broad assessment of theory of mind in adolescence: the complexity of mindreading. *Conscious. Cogn.* 24, 84–97. doi: 10.1016/j.concog.2014.01.003

Breil, C., Kanske, P., Pittig, R., and Böckler, A. (2021). A revised instrument for the assessment of empathy and theory of mind in adolescents: introducing the EmpaToM-Y. *Behav. Res. Methods* 53, 2487–2501. doi: 10.3758/s13428-021-01589-3

Brodsky, J. E., Bergson, Z., Chen, M., Hayward, E. O., Plass, J. L., and Homer, B. D. (2023). Language, ambiguity, and executive functions in adolescents' theory of mind. *Child Dev.* 94, 202–218. doi: 10.1111/cdev.13852

Bronfenbrenner, U. (1979). The ecology of human development. Cambridge: Harvard University Press.

Byrnes, J. P. (2003). "Cognitive development during adolescence" in Blackwell handbook of adolescence. eds. G. R. Adams and M. D. Berzonsky (Oxford, UK: Blackwell Publishing), 227–246.

Cabinio, M., Rossetto, F., Blasi, V., Savazzi, F., Castelli, I., Massaro, D., et al. (2015). Mind-Reading ability and structural connectivity changes in aging. *Front. Psychol.* 6:1808. doi: 10.3389/fpsyg.2015.01808

Caputi, M., and Bosacki, S. (2023). The role of gender in the relation among anxiety, theory of mind, and well-being in early adolescents. *J. Early Adolesc.* 43, 845–866. doi: 10.1177/02724316221130440

Carlo, G., Chavez, F. L. C., and López-Mora, C. (2023). "Prosocial behaviors in adolescence" in The Cambridge handbook of prosociality. In the Cambridge handbook of prosociality: development, mechanisms, promotion. eds. T. Malti and M. Davidov (Cambridge: Cambridge University Press), 163–188.

Carpendale, J. I. M., and Lewis, C. (2004). Constructing an understanding of mind: the development of children's social understanding within social interaction. *Behav. Brain Sci.* 27, 79–151. doi: 10.1017/s0140525x04000032

Carugati, F., and Selleri, P. (2011). "Logica, ragionamento e regole" in Psicologia dell'adolescenza. ed. A. Palmonari (Bologna: Il Mulino), 107–127.

Castelli, I., Massaro, D., and Marchetti, A. (Eds.) (2022). La competenza sociale nei bambini. Valutare e promuovere la Teoria della Mente in età prescolare e scolare. Bergamo: Edizioni Junior.

Castelli, I., Massaro, D., Sanfey, A. G., and Marchetti, A. (2010). Fairness and intentionality in children's decision-making. *Int. Rev. Econ.* 57, 269–288. doi: 10.1007/s12232-010-0101-x

Cavioni, V., Grazzani, I., and Ornaghi, V. (2017). Social and emotional learning for children with learning disability: implications for inclusion. *Int. J. Emot. Educ.* 9, 100–109. doi: 10281/177282

Cefai, C., Camilleri, L., Bartolo, P., Grazzani, I., Cavioni, V., Conte, E., et al. (2022). The effectiveness of a school-based, universal mental health programme in six European countries. *Front. Psychol.* 13:925614. doi: 10.3389/fpsyg.2022.925614

Clifford, M. E., Nguyen, A. J., and Bradshaw, C. P. (2021). Emotion processing associated with aggression in early adolescents: a focus on affective theory of mind. *Aggress. Behav.* 47, 173–182. doi: 10.1002/ab.21936

Colomeischi, A. A., Duca, D. S., Bujor, L., Rusu, P. P., Grazzani, I., and Cavioni, V. (2022). Impact of a school mental health program on Children's and adolescents' socio-emotional skills and psychosocial difficulties. *Children* 9:1661. doi: 10.3390/children9111661

de Villiers, J. G. (2021). "The role(s) of language in theory of mind" in The neural basis of Mentalizing. eds. M. Gilead and K. N. Ochsner (Cham: Springer International Publishing), 423–448.

Devine, R. T., Kovatchev, V., Grumley Traynor, I., Smith, P., and Lee, M. (2023). Machine learning and deep learning systems for automated measurement of "advanced" theory of mind: reliability and validity in children and adolescents. *Psychol. Assess.* 35, 165–177. doi: 10.1037/pas0001186

Devine, R. T., and Lecce, S. (Eds.) (2021). Theory of mind in middle childhood and adolescence: integrating multiple perspectives. New York, NY: Routledge/Taylor & Francis Group.

Dourou, C., Gavriilidou, Z., and Markos, A. (2020). Definitional skills and preferred definition types according to age, gender, educational level and career orientation. *Int. J. Res. Stud. Educ.* 9, 29–49. doi: 10.5861/ijrse.2020.5021

Dumontheil, I., Apperly, I. A., and Blakemore, S. (2010). 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

Fink, E. (2021). "Friendships and theory of mind in middle childhood and adolescence" in Theory of mind in middle childhood and adolescence. eds. R. T. Devine and S. Lecce (New York: Psychology Press), 122–141.

Gabriel, E. T., Oberger, R., Schmoeger, M., Deckert, M., Vockh, S., Auff, E., et al. (2021). Cognitive and affective theory of mind in adolescence: developmental aspects and associated neuropsychological variables. *Psychol. Res.* 85, 533–553. doi: 10.1007/s00426-019-01263-6

Gatta, M., Facca, I., Colombo, E., Svanellini, L., Montagnese, S., and Schiff, S. (2014). Alexithymia, psychopathology and alcohol misuse in adolescence: a population based study on 3556 teenagers. *Neurosci. Med.* 5, 60–71. doi: 10.4236/nm.2014.51009

Hayward, E. O., and Homer, B. D. (2017). Reliability and validity of advanced theory-of-mind measures in middle childhood and adolescence. *Br. J. Dev. Psychol.* 35, 454–462. doi: 10.1111/bjdp.12186

Im-Bolter, N., Agostino, A., and Owens-Jaffray, K. (2016). Theory of mind in middle childhood and early adolescence: different from before? *J. Exp. Child Psychol.* 149, 98–115. doi: 10.1016/j.jecp.2015.12.006

Johnson, C. J., and Anglin, J. M. (1995). Qualitative developments in the content and form of Children's definitions. *J. Speech Lang. Hear. Res.* 38, 612–629. doi: 10.1044/jshr.3803.612

Kinderman, P., Dunbar, R., and Bentall, R. P. (1998). Theory-of-mind deficits and causal attributions. *Br. J. Psychol.* 89, 191–204. doi: 10.1111/j.2044-8295.1998.tb02680.x

Kuhn, D. (2000). "Theory of mind, metacognition and reasoning: a life-span perspective" in Children's reasoning and the mind. eds. P. Mitchell and K. J. Riggs (Hove: Psychology Press), 301–326.

Laghi, F., and Lonigro, A. (2022). "Considerazioni conclusive attraverso uno sguardo retrospettivo: che cosa può dirci la ricerca sulla competenza sociale e sulla Teoria della Mente in adolescenza rispetto alle età precedenti?" in La competenza sociale nei bambini. Valutare e promuovere la Teoria della Mente in età prescolare e scolare. eds. I. Castelli, D. Massaro and A. Marchetti (Bergamo: Edizioni Junior), 387–399.

Lavoie, J., and Talwar, V. (2022). Theory of mind and concealing vs. forthcoming communication in adolescence. *Front. Psychol.* 13:866964. doi: 10.3389/fpsyg.2022.866964

Lebedeva, E. I., Filippou, O. Y., Vilenskaya, G. A., Ulanova, A. Y., and Pavlova, N. S. (2023). Theory of mind and Machiavellianism as factors of intragroup relations in adolescence. *Soc. Psychol. Soc.* 14, 99–117. doi: 10.17759/sps.2023140307

Lecce, S. (2009). La teoria della mente nella vita quotidiana: Il lessico psicologico [Theory of mind within real life context: the mental state talk]. *G. Ital. Psicol.* 36, 21–50. doi: 10.1421/29280

Lecce, S., and Pagnin, A. (2007). Il lessico psicologico: la teoria della mente nella vita quotidiana. Bologna: Il Mulino.

Lerner, R. M., and Steinberg, L. (Eds.) (2009). Handbook of adolescent psychology. Hoboken, New Jersey: Wiley.

Litowitz, B. (1977). Learning to make definitions. *J. Child Lang.* 4, 289–304. doi: 10.1017/S0305000900001665

Lombardi, E., Valle, A., Bianco, F., Castelli, I., Massaro, D., and Marchetti, A. (2022). Supporting mentalizing in primary school children: the effects of thoughts in mind project for children (TiM-C) on metacognition, emotion regulation and theory of mind. *Cogn. Emot.* 36, 975–986. doi: 10.1080/02699931.2022.2067521

Longobardi, E., Spataro, P., and Rossi-Arnaud, C. (2016). Relations between theory of mind, mental state language and social adjustment in primary school children. *Eur. J. Dev. Psychol.* 13, 424–438. doi: 10.1080/17405629.2015.1093930

Marchetti, A., Castelli, I., Valle, A., and Massaro, D. (2016). La teoria della mente nell'invecchiamento: la "conquista della saggezza"? *Ricer. Psicol.* 63–78, 63–78. doi: 10.3280/RIP2016-001005

Marchetti, A., and Cavalli, G. (2013). Le difficoltà emotive nello sviluppo: il caso dell'alessitimia e dell'autolesionismo. Dalla ricerca psicologica e neuroscientifica alla psicoterapia. *Riv. Int. Filos Psicol.* 4, 355–369. doi: 10.4453/rifp.2013.0034

Marchetti, A., and Massaro, D. (2002). Capire la mente: La psicologia ingenua del bambino. Roma: Carocci.

Martinsone, B., Stokenberga, I., Damberga, I., Supe, I., Simões, C., Lebre, P., et al. (2022). Adolescent social emotional skills, resilience and behavioral problems during the COVID-19 pandemic: a longitudinal study in three European countries. *Front. Psych.* 13:942692. doi: 10.3389/fpsyt.2022.942692

Meinhardt-Injac, B., Daum, M. M., and Meinhardt, G. (2020). Theory of mind development from adolescence to adulthood: testing the two-component model. *Br. J. Dev. Psychol.* 38, 289–303. doi: 10.1111/bjdp.12320

Meins, E., Fernyhough, C., Johnson, F., and Lidstone, J. (2006). Mind-mindedness in children: individual differences in internal-state talk in middle childhood. *Br. J. Dev. Psychol.* 24, 181–196. doi: 10.1348/026151005X80174

Miller, S. A. (2022). Advanced theory of mind. New York: Oxford University Press.

Milligan, K., Astington, J. W., and Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Dev.* 78, 622–646. doi: 10.1111/j.1467-8624.2007.01018.x

Moshman, D. (1998). "Cognitive development beyond childhood" in Handbook of child psychology: vol. 2. Cognition, perception, and language. ed. W. Damon (Hoboken, New Jersey: John Wiley & Sons, Inc), 947–978.

Mulvey, K. L., Gönültaş, S., Herry, E., and Strelan, P. (2022). The role of theory of mind, group membership, and apology in intergroup forgiveness among children and adolescents. *J. Exp. Psychol. Gen.* 151, 613–627. doi: 10.1037/xge0001094

Muzi, S. (2020). A narrative review on alexithymia in adolescents with previous adverse experiences placed for adoption, in foster care, or institutions. Prevalence, gender differences, and relations with internalizing and externalizing symptoms. *Mediter. J. Clin. Psychol.* 8, 1–25. doi: 10.6092/2282-1619/mjcp-2449

Nippold, M. A. (1995). School-age children and adolescents. *Lang. Speech Hear. Serv. Sch.* 26, 320–325. doi: 10.1044/0161-1461.2604.320

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., and Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Front. Psychol.* 6:1503. doi: 10.3389/fpsyg.2015.01503

Ornaghi, V., Brockmeier, J., and Grazzani, I. (2014). Enhancing social cognition by training children in emotion understanding: a primary school study. *J. Exp. Child Psychol.* 119, 26–39. doi: 10.1016/j.jecp.2013.10.005

Palmonari, A., and Crocetti, E. (2011). "Identità e concetto di sé" in Psicologia dell'adolescenza. ed. A. Palmonari (Bologna: Il Mulino), 68–90.

Peterson, C. C., and Wellman, H. M. (2019). Longitudinal theory of mind (ToM) development from preschool to adolescence with and without ToM delay. *Child Dev.* 90, 1917–1934. doi: 10.1111/cdev.13064

Piaget, J., and Inhelder, B. (1955). Dalla logica del fanciullo alla logica dell'adolescente., Trad. It. 1971. Firenze: Giunti.

Pinto, G., Primi, C., Tarchi, C., and Bigozzi, L. (2017). Mental state talk structure in Children's narratives: a cluster analysis. *Child Dev. Res.* 2017, 1–7. doi: 10.1155/2017/1725487

Pluck, G., Córdova, M. A., Bock, C., Chalen, I., and Trueba, A. F. (2021). Socio-economic status, executive functions, and theory of mind ability in adolescents: relationships with language ability and cortisol. *Br. J. Dev. Psychol.* 39, 19–38. doi: 10.1111/bjdp.12354

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/S0140525X00076512

Sebastian, C. L. (2015). Social cognition in adolescence: social rejection and theory of mind. *Psicol. Educ.* 21, 125–131. doi: 10.1016/j.pse.2015.08.004

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

Serafin, M., and Surian, L. (2004). Il test degli occhi: uno strumento per valutare la "Teoria della Mente.". *G. Ital. Psicol.* 4, 839–862. doi: 10.1421/18849

Siegal, M., and Surian, L. (Eds.) (2011). Access to language and cognitive development. New York: Oxford University Press.

Snow, C. E. (1990). The development of definitional skill. *J. Child Lang.* 17, 697–710. doi: 10.1017/S0305000900010953

Snow, C., Cancino, H., Gonzales, P., and Shriberg, E. (1989). "Giving formal definitions: an oral correlate of school literacy" in Literacy in classrooms. ed. D. Bloome (Norwood: Ablex), 233–249.

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends Cogn. Sci.* 9, 69–74. doi: 10.1016/j.tics.2004.12.005

Stępień-Nycz, M., Bosacki, S., and Białecka-Pikul, M. (2021). Theory of mind in adolescence: developmental change and gender differences. *J. Early Adolesc.* 41, 424–436. doi: 10.1177/0272431620919164

Symeonidou, I., Dumontheil, I., Chow, W.-Y., and Breheny, R. (2016). Development of online use of theory of mind during adolescence: an eye-tracking study. *J. Exp. Child Psychol.* 149, 81–97. doi: 10.1016/j.jecp.2015.11.007

The Jamovi Project (2022). jamovi [Computer Software]. Available at: https://www.jamovi.org.

Tolchinsky, L., and Berman, R. A. (2023). Growing into language. Oxford: Oxford University Press.

Valle, A., Massaro, D., Castelli, I., and Marchetti, A. (2015). Theory of mind development in adolescence and early adulthood: the growing complexity of recursive thinking ability. *Eur. J. Psychol.* 11, 112–124. doi: 10.5964/ejop.v11i1.829

Van Den Bos, E., De Rooij, M., Sumter, S. R., and Westenberg, P. M. (2016). Continued development of recursive thinking in adolescence: longitudinal analyses with a revised recursive thinking test. *Cogn. Dev.* 37, 28–41. doi: 10.1016/j.cogdev.2015.11.002

Vetter, N. C., Weigelt, S., Döhnel, K., Smolka, M. N., and Kliegel, M. (2014). Ongoing neural development of affective theory of mind in adolescence. *Soc. Cogn. Affect. Neurosci.* 9, 1022–1029. doi: 10.1093/scan/nst081

Waite-Jones, J. M., and Rodriguez, A. M. (2022). Psychosocial approaches to child and adolescent health and wellbeing. Cham: Springer International Publishing.

Wang, X., Su, Y., Pei, M., and Hong, M. (2021). How self-other control determines individual differences in adolescents' theory of mind. *Cogn. Dev.* 57:101007. doi: 10.1016/j.cogdev.2021.101007

Watson, R. (1985). Towards a theory of definition. *J. Child Lang.* 12, 181–197. doi: 10.1017/S0305000900006309

Watson, R. (1995). Relevance and definition. *J. Child Lang.* 22, 211–222. doi: 10.1017/S0305000900009703

Wellman, H. M. (2012). Theory of mind: better methods, clearer findings, more development. *Eur. J. Dev. Psychol.* 9, 313–330. doi: 10.1080/17405629.2012.680297

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128. doi: 10.1016/0010-0277(83)90004-5

World Medical Association (2013). World medical association declaration of Helsinki. *JAMA* 310:2191. doi: 10.1001/jama.2013.281053

Zerwas, S., Balaraman, G., and Brownell, C. (2004). Constructing an understanding of mind with peers. *Behav. Brain Sci.* 27:130. doi: 10.1017/S0140525X04590039

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 17 October 2024
DOI 10.3389/fpsyg.2024.1461980

OPEN ACCESS

EDITED BY
Simone G. Shamay-Tsoory,
University of Haifa, Israel

REVIEWED BY
Valeria Blasi,
Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy
Antonio Narzisi,
Stella Maris Foundation (IRCCS), Italy

*CORRESPONDENCE
Roberta Fadda
✉ robfadda@unica.it

RECEIVED 09 July 2024
ACCEPTED 30 September 2024
PUBLISHED 17 October 2024

CITATION
Fadda R, Congiu S, Doneddu G, Carta M, Piras F, Gabbatore I and Bosco FM (2024) Th.o.m.a.s.: new insights into theory of mind in adolescents with autism spectrum disorder. *Front. Psychol.* 15:1461980.
doi: 10.3389/fpsyg.2024.1461980

Th.o.m.a.s.: new insights into theory of mind in adolescents with autism spectrum disorder

Roberta Fadda[1]*, Sara Congiu[1], Giuseppe Doneddu[2], Marinella Carta[3], Francesco Piras[3], Ilaria Gabbatore[4,5] and Francesca M. Bosco[4]

[1]Department of Pedagogy, Psychology, Philosophy, University of Cagliari, Cagliari, Italy, [2]Center for Autism and Related Disorder, Nuovo Centro Fisioterapico Sardo, Cagliari, Italy, [3]Center for Pervasive Developmental Disorder, AO Brotzu, Cagliari, Italy, [4]Department of Psychology, GIPSI Research Group, University of Turin, Turin, Italy, [5]Department of Humanities, University of Turin, Turin, Italy

Previous studies indicated atypical Theory of Mind (ToM) abilities in individuals with autism spectrum disorder (ASD) at different ages. However, research focused on adolescents with ASD is still rare. This study aims to fill the gaps in the literature, by investigating ToM abilities in adolescents with ASD and in a group of typically developing ones. We applied the Theory of Mind Assessment Scale (Th.o.m.a.s.), a semi-structured interview that allows a multi-dimensional measurement of ToM, including different perspectives (first/s-order, first/third-person, egocentric/allocentric), various mental states (emotions, desires, beliefs) and metacognitive abilities related with mental states (awareness, relation, and strategies). The results indicated that ToM develops atypically in ASD, with strengths and weaknesses. First, participants with ASD were comparable to controls in some specific ToM aspects, i.e., third-person ToM, both from an egocentric and an allocentric perspective. However, they were significantly weaker in attributing an understanding of the mental states of others, both in first- and second-order ToM scenarios. Second, they showed the same level of awareness about mental states as controls, but they were significantly weaker in conceptualizing the relationship between mental states and behavior. Also, they found it very difficult to think about possible strategies that they or others might employ to realize desires and needs. Finally, they performed similarly to controls in understanding emotions, while they poorly understood desires and beliefs. These results point out the distinctive characteristics of ToM development in individuals with ASD, with important implications for individualized interventions.

KEYWORDS

autism spectrum disorder, theory of mind, mindreading, assessment, clinical interview, adolescents

1 Introduction

Autism spectrum disorder (ASD) is characterized by persistent deficits in social communication and social interaction across multiple contexts. Specifically, individuals with ASD might show deficits in social–emotional reciprocity, and reduced sharing of interests, emotions, or affect with others. They might also show a deficit in nonverbal communicative behaviors used for social interaction and difficulties in developing, maintaining, and understanding relationships. Moreover, ASD is often characterized by restricted repetitive patterns of behavior (American Psychiatric Association, 2013).

The roots of the social and communicative deficit in ASD can be identified at an early age when individuals with ASD fail to develop joint attention abilities, which allow for representing the same focus of interest of another person. The lack of joint attention abilities, early in infancy, is believed to exert a cascade effect on the subsequent poor development of Theory of Mind (ToM) abilities in childhood (Mundy, 2018; Congiu et al., 2016). ToM is the ability to attribute mental states like desires, emotions, intentions, and beliefs to oneself and to others to explain and predict behavior (Wimmer and Perner, 1983). An atypical acquisition of ToM abilities is supposed to hamper social interactions in individuals with ASD during their whole lifespan (Andreou and Skrimpa, 2020; Brewer et al., 2017; Angeleri et al., 2016).

Research about ToM in ASD has mostly involved children. Recently, the interest in investigating ToM in older individuals with ASD is growing (Livingston et al., 2019). According to a recent meta-analysis (Gao et al., 2023), which considered 110 studies including 3,205 participants with ASD and 3,675 typically developing (TD) controls (mean age ≥ 18 years), indicated that late adolescents (18–24 years old) and adults with ASD demonstrate a weak performance in ToM task as compared to controls. According to Gao et al. (2023), ToM tasks that have been widely used in older individuals with ASD can be classified into four categories: *reading comprehension, perceptual scene comprehension, comprehensive scene comprehension,* and *self–other processing.*

The *reading comprehension* tasks, like for example the *Strange Stories* test (Happé, 1994), require participants to infer a character's mental state and subsequent behavior based on the reading of relevant information in verbal vignettes.

The *perceptual scene comprehension* tasks, like for example the *animation task* (Abell et al., 2000), evaluate the ability to infer mental states behind the movement of geometrical forms (i.e., triangles) without any explicit language information and in simple social scenarios.

The *comprehensive scene comprehension* tasks, like for example *the Strange Stories Film task* (Murray et al., 2017), which consists of video scenarios based on the original *Strange Stories* (Happé, 1994): irony, double bluff, pretense, joke, appearance/reality, white-lie, persuasion, misunderstanding, forgetting, contrary emotions, and idioms. It tests the ability to attribute mental states to the characters displayed in the videos. Finally, the *self–other processing* tasks require processing a conflict between one's own and others' mental states and responding by shifting between one's own and others' points of view (Deschrijver and Palmer, 2020), which comprises explicit and implicit versions. The explicit tasks depend on language processing to stimulate individuals' inferences about mental states, like for example the *Sandbox task* (Sommerville et al., 2013). The implicit tasks elicit rapid mental state attribution, independently from language. Implicit tasks include for example the *Reading the Mind in the Eyes test* (Baron-Cohen et al., 1997, 2001), which requires matching images of the eyes with mental state labels. Another example of implicit tasks are the eye-tracking measures of participants' visual attention while observing an agent who holds a false belief (e.g., Senju et al., 2009).

The results of the meta-analysis (Gao et al., 2023) indicated a significant moderating effect of the ToM task's type, since the ToM difference in *reading comprehension tasks* and *comprehensive scene comprehension tasks* was larger than that in *perceptual scene comprehension tasks* and in *self–other processing tasks*. This means that adolescents and adults with ASD might display different ToM competencies, depending on the tasks. Moreover, the ToM tasks used so far have been basically shaped by the ones originally used for young typically developing children (e.g., Baillargeon et al., 2010; Bowler, 1992; Wimmer and Perner, 1983), which tend to reduce mindreading abilities in terms of a presence/absence phenomenon (Livingston et al., 2019). Passing these kinds of tasks might not reflect the actual ToM abilities of older individuals with ASD, masking the possible difficulties that they might still experience in thinking about mental states. Indeed, ToM has a complex nature that cannot be reduced to an on–off or an all-or-nothing functioning (Tirassa et al., 2006). It is based on a developmental progression of a variety of insights about mental states like intentions, emotions, desires, knowledge, and beliefs (see, e.g., Wellman, 2014). It includes different dimensions, like the understanding of the first- and third-person perspective, which is mediated by different processes, and it recruits several types of knowledge (Nichols and Stich, 2004). It also includes the distinction between an egocentric perspective, in which the others are represented in relation to the self, and an allocentric one, in which others' mental states are represented independently from the self (Frith and de Vignemont, 2005).

In general, previous studies using classical ToM tasks did not investigate the richness of the actual mentalization abilities in adolescent and adult individuals with ASD. This leaves open the question of whether ToM abilities of subjects in this age range who pass the classical ToM tasks are comparable to those displayed by age-matched TD controls. It is possible that other compensatory abilities, which are known to be functional to pass the classical ToM tasks, might lead individuals with ASD to interpret others' behaviors in a very concrete and logical way, by considering external events to cause others' behavior without the mediating effect of mental representations.

Also, it is important to consider that with aging, social-, verbal-, and nonverbal abilities tend to develop in ASD adults (Howlin and Magiati, 2017; Ratto and Mesibov, 2015). The development of these abilities might make up for the atypical ToM to a certain extent. It is possible that some adults with high-functioning ASD develop some cognitive compensation strategies that allow them to effectively perform ToM tasks thanks to their general cognitive and language skills (Frith, 1994; Begeer et al., 2010), bypassing the problem of a lack of ToM abilities. Since qualitative difficulties in social interaction persist for these individuals in everyday life, scientists assume that the use of compensatory strategies leads to passing some experimental ToM tasks (Senju et al., 2009). In line with this hypothesis, several studies indicated that linguistic and cognitive abilities, as well as executive control, significantly affect the performance of adolescents with ASD in succeeding in classical ToM tasks. A recent study explored the possible association between ToM, Executive Functioning (EF), and parent-reported measures of social communication and restricted and repetitive behaviors (RRBs) in adolescents with ASD (Jones et al., 2018). A sample of 100 adolescents with ASD (mean age 15 years 6 months) was tested by a series of ToM tasks: a false belief task, the *Strange Stories*, the Frith-Happé *animation task*, the *Reading the mind in the eyes task*. A structural equation modeling was used to verify the possible associations between ToM abilities, EF, and parent-reported measures of social communication and restricted and repetitive behaviors (RRBs). The results indicated that ToM abilities were associated with both social communication symptoms and RRBs. EF was a correlate of ToM but had no direct association with parent-reported symptom expression.

Also, according to the weak central coherence theory, adults with ASD exhibit a detail-focused style of cognition (Frith and Happé, 1994), potentially affecting information processing in ToM tasks. Furthermore, individuals with ASD might encounter difficulties in cognitive flexibility and inhibition control (Hill, 2004), which may lead to difficulties in shifting and controlling in perspective between self and others (Devine and Lecce, 2021).

In summary, previous studies indicated that individuals with ASD develop atypical ToM abilities in adolescence and adulthood. However, this research topic still deserves further investigation. On one hand, most ToM research has been focused on children, informing our understanding of mentalistic abilities and related atypical social behavior during childhood. On the other hand, only a small part of ToM research focused on adolescents and adults with ASD, therefore there is still a poor understanding of their actual mentalization abilities. Also, there are still a series of methodological concerns about the sensitivity of ToM tasks used so far. These tasks consider ToM like an all-or-nothing function, thus hampering the possibility to detect subtle distinctive features of ToM abilities in autism that might be important to develop effective intervention programs. Moreover, classical ToM tasks tap different processes underlying ToM abilities, like linguistic, cognitive abilities, and executive functioning, that might compensate for possible mindreading difficulties, leading to pass classical laboratory tasks. Finally, none of these tasks have been standardized not only in TD controls but also in other clinical populations, leaving open the question of whether some ToM deficits might be distinctive of ASD or not. These methodological concerns call for more sensitive tools to investigate mentalistic abilities in older individuals with ASD.

In this study, we investigated ToM abilities in adolescents with ASD by applying a multidimensional conceptualization of ToM abilities (Bosco et al., 2009b), compared with a group of typically developing matched controls, with the Theory of Mind Assessment Scale (Th.o.m.a.s.).

Th.o.m.a.s is a semi-structured interview that allows a multi-component and ecological measurement of different dimensions of ToM (Bosco et al., 2006): egocentric vs. allocentric perspective; beliefs vs. desires vs. positive emotions vs. negative emotions; awareness (the ability to perceive and differentiate mental states in oneself and in others) vs. causal relationships between mental states and behavior vs. efficient strategies to achieve desired states. The scale has been standardized in typical populations of adolescents and adults (Bosco et al., 2014; Bosco et al., 2016).

By adopting a multidimensional approach to investigate different ToM aspects, we aim to provide a complete, detailed, and comparable profile of mentalizing abilities in adolescents with ASD, in which specific components or sub-skills might be less or more impaired than others. We predicted that individuals with ASD, in line with other clinical populations, would show generally lower ToM abilities than controls, especially in high-level mental states like beliefs and second-order perspective.

2 Methods

2.1 Participants

We enrolled 20 participants with ASD in this study. One participant was excluded because he had a history of cognitive delay. Another one was excluded because he had a chronological age well beyond the age range of young adulthood (45 yrs). The final sample included 18 participants with ASD (3 Females; 15 Males), mean age 16 years and 5 months (± 3), and 18 typically developing controls (3 Females; 15 Males), mean age 16 years and 3 months (± 3). All participants with ASD had been diagnosed by expert clinicians and fulfilled the international diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders 5th edition, DSM-5 (American Psychiatric Association, 2013). The diagnosis has been confirmed with the Autism Diagnostic Observation Schedule at the time of onset (Lord et al., 2005, 2013). Only 12 participants received a re-evaluation for symptom severity in adolescents according to age and verbal fluency: 8 participants were evaluated with module 4 (mean score of communication + social interaction = 6,666, SD ±3.605) while 4 were evaluated with module 3 (mean score of communication + social interaction = 7, SD ±2.966). All the participants used phrases with more than five words. Participants received education in mainstream classes in regular middle or high school. The full-scale IQ (M = 108.69, SD = 14.323, Range = 80–141) was estimated using the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2012). The IQ of two participants, which were, respectively, 89 and 123, was evaluated with the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV; Wechsler, 2013). All the subjects were recruited through the Center for Pervasive Developmental Disorder of Azienda Ospedaliera Brotzu, in Cagliari, Italy.

2.2 Materials and procedures

Th.o.m.a.s. is a semi-structured interview to investigate ToM (Bosco et al., 2006; Bosco et al., 2016), and has proven effective in a number of clinical populations (e.g., Bosco et al., 2009a; Laghi et al., 2014; Colle et al., 2019). It includes 37 open-ended questions that ask participants to express their understanding of their own and others' mental states. The questions are organized into four scales: Scale A, I–Me (that investigates the interviewee's knowledge of her own mental states—1st person ToM in an egocentric perspective); Scale B, Other–Self (which investigates the knowledge that, according to the interviewee, the other persons have of their own mental states, independently of the subject's perspective—3rd person ToM in an allocentric perspective). Scale C, I–Other (which investigates the interviewee's knowledge of the mental states of other persons 3rd person ToM in an egocentric perspective); Scale D, Other–Me (which investigates the knowledge that, from the interviewee's point of view, the others have of her mental states comparable to a 2nd order ToM—in an allocentric perspective). Each scale is divided into three subscales that, respectively, explore the dimensions of Awareness (the interviewee's ability to perceive and differentiate mental states in herself and in others), Relation (the interviewee's ability to recognize causal relations between different mental states and between them and the resulting behaviors), and Realization (the interviewee's ability to adopt effective strategies to achieve a desired state). The interview also allows to focus on the interviewee's perspectives on epistemic states (knowledge, beliefs and so on), volitional states (desires, intentions, and so on), and positive and negative emotions.

To evaluate participants' general Theory of Mind abilities, we administered a classical ToM task, consisting of a selection of four Strange Stories (Happé, 1994) to both groups of participants. Each participant was tested individually in a quiet room after signing a written consent and parents signed the consent for participants under 18 years of age. All Th.o.m.a.s. interviews were audio-recorded and then transcribed. The transcriptions were rated by two independent judges, who were blind to whether participants belonged to the experimental or the control group. Each judge was asked to evaluate each answer with a score from 0 to 4, according to the given rating criteria. To assess the inter-rater agreement an Intraclass Correlation Coefficient (ICC) was calculated on the 30% of the sample. The ICC was 0.865, indicating substantial reliability (Shrout, 1998). The study was approved by the Institutional Review Board committee of the Department of Pedagogy, Psychology, Philosophy of the University of Cagliari (Italy).

3 Results

Preliminary, we investigated participants' general Theory of Mind abilities in the Strange Stories (Happé, 1994). The results indicated that participants with ASD were as able as controls in attributing mental states to the characters of the Strange Stories ($t = 1.926$; df = 33; $p = 0.063$).

To compare the performance of individuals with autism and typically developing controls on the Th.o.m.a.s. scales, we performed a repeated measures ANOVA with a two-level between-subjects factor (ASD vs. control group) and a four-level within-subjects factor (Th.o.m.a.s. scales: A (I-Me), B (Other-Self), C (I–Other), and D (Other–Me)).

The analysis revealed an effect of the group ($F_{(1,34)} = 10.47$; $p = 0.003$; $\eta^2_p = 0.235$), an effect of the scale ($F_{(3,102)} = 17.127$; $p < 0.001$; $\eta^2_p = 0.335$) and a significant Scale×Group interaction ($F_{(3,102)} = 3.871$; $p = 0.011$; $\eta^2_p = 0.102$).

To better explore such a result, we ran a series of t-tests (Bonferroni correction for multiple comparisons: alpha ≤0.012), which revealed that the performance of the ASD group was significantly lower than that of the control group on scale C (I-Other), investigating 3rd person ToM in an egocentric perspective ($t_{(34)} = 3.462$; $p = 0.001$; d = 0.599) and D (other-me), investigating egocentric second order ToM ($t_{(34)} = 4.075$; $p < 0.001$; d = 0.722) while no significant differences were detected in the performance of subjects with ASD and controls on scales A (I-Me), investigating first person ToM ($t_{(34)} = 1.708$; $p = 0.097$; d = 0.294) and B (Other-Self) investigating 3rd person ToM from an allocentric perspective ($t_{(34)} = 1.923$; $p = 0.063$; d = 0.350) (Figure 1).

In order to compare the performance of the two groups on Awareness, Relation, and Realization, we also run a repeated measures ANOVA with a two-level between-subjects factor (ASD vs. control group) and a three-level within-subjects factor (Th.o.m.a.s. Subscales: Awareness, Relation, and Realization). The analysis showed an effect of the group ($F_{(1,34)} = 10.687$; $p = 0.002$; $\eta^2_p = 0.239$), an effect of the subscale ($F_{(2,68)} = 3.976$; $p = 0.023$; $\eta^2_p = 0.105$) and a significant Group×Subscale interaction ($F_{(2,68)} = 3.124$; $p = 0.05$; $\eta^2_p = 0.084$).

To explore further these results, we ran a series of *t*-tests (Bonferroni correction for multiple comparisons: alpha≤0.017). As shown in Figure 2, the performance of the ASD group was significantly lower than that of the control group on the Causal Relation ($t_{(34)} = 2.783$; $p = 0.009$; d = 0.476) and Realization subscales ($t_{(34)} = 3.948$; $p < 0.001$; d = 0.639) but not on the Awareness subscale ($t_{(34)} = 2.284$; $p = 0.029$; d = 0.361).

In order to investigate the performance of the two groups in positive emotions, negative emotions, desires, and beliefs, we run a repeated measures ANOVA with a two-level between-subjects factor (ASD vs. control group) and a four-level within-subjects factor

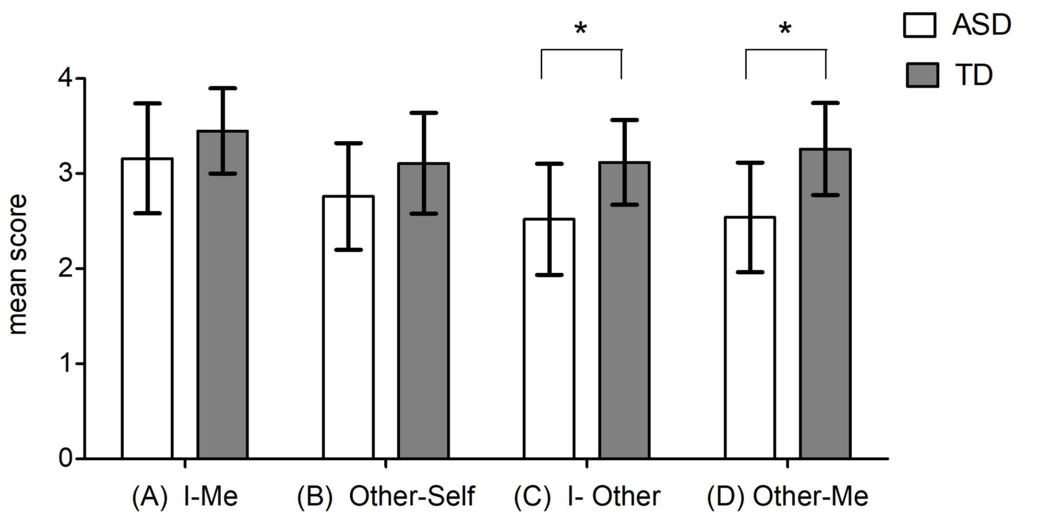


FIGURE 1
ASD vs. controls: mean scores at Th.o.m.a.s. scales: Scale A, I–Me (that investigates the interviewee's knowledge of her own mental states—1st person ToM in an egocentric perspective); Scale B, Other–Self (which investigates the knowledge that, according to the interviewee, the other persons have of their own mental states, independently of the subject's perspective—3rd person ToM in an allocentric perspective); Scale C, I–Other (which investigates the interviewee's knowledge of the mental states of other persons 3rd person ToM in an egocentric perspective); Scale D, Other–Me (which investigates the knowledge that, from the interviewee's point of view, the others have of her mental states comparable to a 2nd order ToM—in an allocentric perspective). Error bars depict a 95% confidence interval. *$p < 0.012$, generated by independent *t*-test.

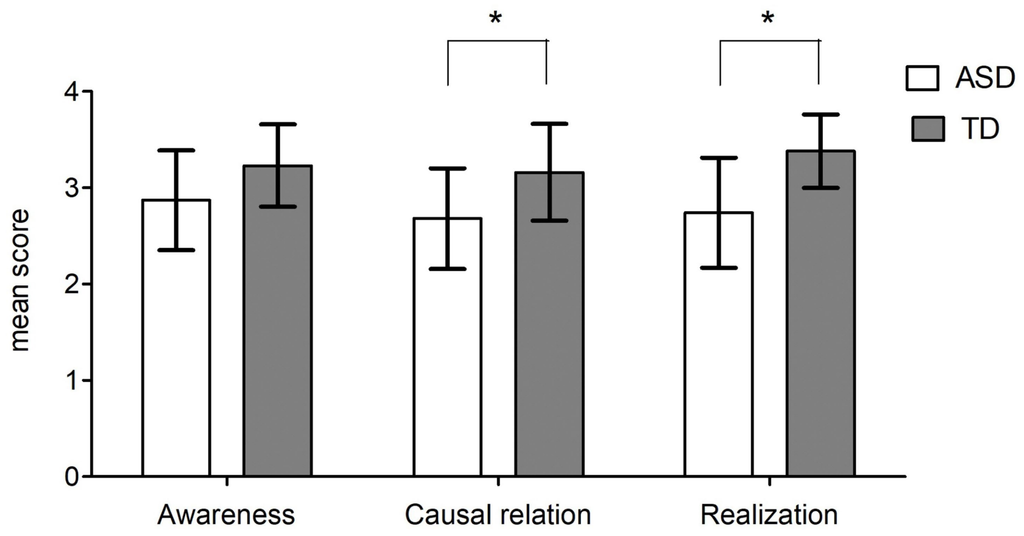


FIGURE 2
ASD vs. controls: mean scores at Th.o.m.a.s. subscales: Awareness (the interviewee's ability to perceive and differentiate mental states in herself and in others), Relation (the interviewee's ability to recognize causal relations between different mental states and between them and the resulting behaviors), and Realization (the interviewee's ability to adopt effective strategies to achieve a desired state). Error bars depict a 95% confidence interval. $*p < 0.017$, generated by independent *t*-test.

FIGURE 3
ASD vs. controls: mean scores at Th.o.m.a.s. dimensions: positive emotions (+ emotions); negative emotions (−emotions); desires; beliefs. $*p < 0.017$, generated by independent *t*-test. Error bars depict a 95% confidence interval.

(Th.o.m.a.s. Dimensions: positive emotions, negative emotions, desires, and beliefs). The analysis revealed a significant group effect ($F_{(1,34)} = 10.990$; $p = 0.002$; $\eta^2_p = 0.244$), a significant effect of dimension ($F_{(3,102)} = 7.058$; $p < 0.001$; $\eta^2_p = 0.172$) and a significant Group×Dimension interaction ($F_{(3,102)} = 4.029$; $p = 0.009$; $\eta^2_p = 0.106$). As shown in Figure 3, a series of t-tests (Bonferroni correction for multiple comparisons: alpha≤0.012) indicated that the performance of the ASD group was significantly lower than that of the control group on Desires ($t_{(34)} = 3.815$; $p = 0.001$; $d = 0.594$) and Beliefs ($t_{(34)} = 3.824$; $p = 0.001$; $d = 0.708$) but not on the Positive emotions ($t_{(34)} = 2.136$; $p = 0.04$; $d = 0.379$) and Negative emotions ($t_{(34)} = 2.033$; $p = 0.05$; $d = 0.326$).

Finally, we analyzed adolescents with ASD's performance within-group in the four scales of the Th.o.m.a.s, by running an ANOVA, with four levels on within-subjects factors (scale type: A, I–Me; B, Other–Self; C, Me–Other; D, Other–Me). We found a significant effect of the type of scale on the ASD's mean scores ($F_{(3,51)} = 13.606$, $p < 0.001$, $\eta^2_p = 0.445$). Specifically, *post hoc* pairwise comparison Bonferroni revealed that participants with ASD scored higher at scale A (I–Me), which assesses first-person ToM than at all the other three scales: B (Other–Self) ($p = 0.004$) and C (Me–Other) ($p = 0.001$), both of which assess third-person ToM, and D (Other–Me) ($p < 0.001$), which assesses ToM with a second-level inference. No significant differences existed between the latter three scales (Table 1).

TABLE 1 Comparison between the four scales of the Th.o.m.a.s with an ANOVA within-group in participants with autism spectrum disorder: Scale A indicates I-Me; Scale B indicates Other-Self; Scale C indicates Me-Other; Scale D indicates Other-Me.

Scale	Mean	SD	F statistics; significance p, η^2
Scale A	3.16	0.577	$F_{(3,51)} = 13.606$, $p < 0.001$, $\eta^2_p = 0.445$
Scale B	2.76	0.562	
Scale C	2.52	0.583	
Scale D	2.54	0.576	

TABLE 2 Comparison between the three subscales of the Th.o.m.a.s with an ANOVA within-group in participants with autism spectrum disorder: Awareness (the ability to understand self and/or other mental states), Relation (the understanding of the relationship between mental states and behavior) and Realization (the strategies that a person can use to modify self or others' mental states).

Scale	Mean	SD	F statistics; significance p, η^2
Awareness	2.87	0.517	$F_{(1,17)} = 2.004$, $p = 0.175$, $\eta^2_p = 0.105$
Relation	2.68	0.523	
Realization	2.74	0.571	

TABLE 3 Comparison between the four dimensions of the Th.o.m.a.s with an ANOVA within-group in participants with autism spectrum disorder: Beliefs, Desires, Positive emotions, Negative emotions.

Scale	Mean	SD	F statistics; significance p, η^2
Beliefs	2.54	0.596	$F_{(3,51)} = 10.250$, $p < 0.001$, $\eta^2_p = 0.376$
Desires	2.66	0.487	
Positive emotions	2.79	0.561	
Negative emotions	3.01	0.519	

We also conducted a ANOVA to investigate ASD's performance at the Th.o.m.a.s. subscales with three levels within-subjects factor (subscale type: Awareness, Relation, Realization). As it is possible to see in Table 2, there wasn't an effect of the type of subscale and thus no difference between ASD's mean performance ($F_{(1,17)} = 2.004$, $p = 0.175$, $\eta^2_p = 0.105$).

We investigated ASD's performance at the four Th.o.m.a.s. dimensions (Beliefs, Desires, Positive emotions, and Negative emotions). As shown in Table 3, we found a significant effect ($F_{(3,51)} = 10.250$, $p < 0.001$, $\eta^2_p = 0.376$). *Post hoc* pairwise comparisons revealed that participants with ASD scored higher on Negative emotions than on Positive emotions ($p = 0.046$), Desires ($p = 1$), and Beliefs ($p = 0.078$). No significant differences existed between the latter three scales.

4 Discussion

Our study investigated ToM abilities in adolescents with ASD with a multidimensional approach. The results indicated that ToM develops atypically in ASD, with weakness in some dimensions but not in all. First, participants with ASD scored significantly weaker than controls in the egocentric perspective. They scored also weaker than controls in the second-order understanding of mental states. However, their performance scores were comparable to those of controls when they had to reflect on mental states referred to self, from the self (egocentric perspective), and when they had to reflect on mental states that others refer to themselves (allocentric perspective).

Second, even though participants with ASD were as good as controls in the awareness of the different mental states considered in the interview, they showed significant difficulty in conceptualizing the relationship between mental states and behaviors and the possible strategies to realize desires and needs. Thus, individuals with ASD seem to be characterized by a rather descriptive Theory of Mind but not by an explanatory one. Being unable to connect different types of mental states with perceptions and actions might hamper their ability to use the knowledge about mental states to successfully affect others' mental states. This resembles the difference between declarative and procedural knowledge about the world stored in long-term memory, which does not always match. It is possible that individuals with autism spectrum disorder might learn what a mental state actually is in a descriptive way but that they lack the possibility, for various reasons, to put into practice their knowledge. Interpreting these results in terms of cognitive models of memory, it is like if in persons with ASD the *episodic buffer* of the working memory (Baddeley, 2000) would not adequately support the memorization of the procedures needed to act adaptively in the social world. Individuals with ASD might often be more sensitive to non-social rather than to social information about the real world. The first might therefore end up being stored in their *episodic* and *semantic memory* at the expense of the second. This hypothesis is in line with the idea that ToM deficit is not the only model to explain social deficit in ASD, but that also sensory and perceptual frameworks provide an alternative explanation (e.g., Garfinkel et al., 2016).

When we considered what mental states individuals with ASD are particularly aware of, we found that participants with ASD were comparable to controls in understanding emotions, while their performance was significantly lower in understanding other mental states like desires and beliefs. Interestingly a within-group analysis revealed that they were particularly able to understand negative emotions. These results might indicate that individuals with ASD, despite well-developed linguistic and cognitive abilities, might still show a delay in the development of Theory of Mind (Yu and Wellman, 2023). Several studies in typically developing children indicated that ToM evolves with age, from infancy to childhood (see, e.g., Wellman and Liu, 2004). The understanding of emotions is the first to appear at around 2 years, followed by the understanding of desires and true beliefs at around 3 years of age, and finally false beliefs at around 4 years. It is like individuals with ASD, even though they might have a mental age equal to their peers during adolescence, are still immature in terms of their reasoning about mental states. They achieve the same level of knowledge about emotions, in particular the negative ones, which is basically the first step of the development in ToM. Individuals who interpret human behavior mainly in terms of emotions might be strongly dependent on reality, missing the constructivist activity of the mind. Thus, a certain state of the world necessarily determines specific emotions but does not elicit a desire or a belief.

ToM functioning in individuals with autism spectrum disorders shows interesting similarities and differences compared to other clinical conditions, which are characterized by significant difficulties in social relationships. In general, individuals with ASD are able to think about their own mental states from a first-order perspective, as measured in Scale A (I-Me). This is a strength also in other clinical populations, like individuals with schizophrenia, eating disorders, and borderline personality disorder (e.g., Bosco et al., 2009a; Laghi et al., 2014). However, we found that our participants with ASD show significant difficulties in recursive thinking, which is necessary to represent second-order mental states, as indicated in Scale D (Other-Me). This is in common with the other clinical conditions previously mentioned, indicating a generalized disruption of the ability to conceive the constructivist nature of other people's minds, which can go far beyond the objective world (Bosco et al., 2009a; Laghi et al., 2014).

Although these similarities, autism spectrum disorder seems to be a peculiar condition, rather different from other clinical populations. Schizophrenia, which is a psychiatric disorder, negatively affects all the dimensions of Theory of Mind, leading to a severe misinterpretation of the social world (Bosco et al., 2009a). Mental Disorders, like eating disorders (Laghi et al., 2014) and borderline personality disorders (Colle et al., 2019), lead to hypermentalization, which is the tendency to base one's own interpretation of social behaviors upon the content of the mind of others rather than on objective observable data. So that the interpretation of the social world might be inaccurate but only with respect to Scale B (Other-Self), which targets the allocentric third-person perspective (Colle et al., 2019).

Indeed, autism spectrum disorder, which is a long-life neurodevelopmental disorder, seems to be characterized by an inaccurate egocentric third-person perspective, as indicated by Scale C (I-Other). They consider another person's mental states as extremely independent from themselves. Other minds are conceived as deeply opaque and highly unpredictable (i.e., "*How can I know what he feels if he does not tell me?*"). Also, they do not know how to influence others' mental states through their behavior. Individuals with ASD do not hypermentalize. In general, they do not rely on interpersonal expectations. This means that they elaborate their interpretation of social behaviors upon their own intrapersonal expectations, which are grounded on their own state of knowledge about objective facts, as expressed in sentences like "My friend is my same age so he must hold the same desires as me, like, for example, getting a good grade at school." Also, they base their interpretation of other minds on learned social rules and cliché, which might be a compensatory strategy to adapt to the social world. It is possible that their difficulties in representing other perspectives might lead them to hyper-generalize prototypical situations associated with a specific mental state rather than develop an effective ToM, based on their own experiences. Also, the explanation might be the other way around. It is possible that their social deficit might expose them to a major risk of isolation compared to other people. Thus, they might lack the opportunities to live the typical teenage or young adult experiences which might be fundamental to inform the development of an effective ToM. This explanation is in line with their tendency to identify their desires with the possession of items or with activities that are more typical for younger ages. Also, participants with ASD showed better performances in negative emotions compared to positive emotions, desires and beliefs. These results are in line with TD adolescents (Bosco et al., 2016) and with non-suicidal self-injury (NSSI) adolescents (Laghi et al., 2016). It seems that adolescence in general is characterized by trouble and existential confusion, which might induce individuals to be more focused on their negative emotions.

There are some possible limitations of the study that need to be acknowledged. This task requires participants to speculate on their own or others' mental states and subsequent behavior based on memory-stored information. Thus, participants are required to retrieve prototypical information from their own experiences to infer mental states. So, information stored in long-term and working memory might play a central role. Also, this type of task also reflects linguistic skills, which are known to be related to effective ToM reasoning (Livingston and Happé, 2017). This means that our results cannot be generalized to the entire autistic spectrum. Another limitation that needs to be addressed is the reduced size of the sample. Moreover, Theory of Mind abilities might be sensitive to individual differences in symptom severity. Even though our participants were all fluent in language, received education in mainstream classes, and had the cognitive resources to attend the interview the information about the ADOS was unfortunately incomplete so we could not use it in the analysis.

As a possible future direction, we do believe that a developmental perspective might help to account for the different advances that occur in childhood and continue into adulthood in individuals with ASD. Moreover, it is important to continue to study ToM abilities in the lifespan with longitudinal studies, from adolescents to adulthood and to the elderly age, in larger samples of subjects.

5 Conclusion

Theory of mind is a progression of understandings about mental states, some of which may be less developed in individuals with autism at a certain point in life. Also, autism presents as a spectrum, thus some individuals with ASD achieve more theory-of-mind insights, and some achieve less. It is important to acknowledge that both theory of mind and autism are developmental phenomena, in which some advances occur in childhood and continue into adulthood. Thus, a functional and dynamic evaluation of Theory of Mind might allow us to understand that individuals with autism do not lack theory of mind overall, as a static and core characteristic. Instead, they can come up to develop many theory-of-mind competencies, although on a delayed timetable (Loukusa et al., 2023; Yu and Wellman, 2023).

Intervention in adolescents should focus mainly on second-order representation (recursive thinking) and on third-person allocentric perspective, which seems to be a long-lasting deficit in this population. Also, participation in real-life social experiences in various contexts should be recommended at this age, to promote procedural knowledge about the relationship between other high-level mental states and behavior, like beliefs and false beliefs.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethic Committee of the University of Cagliari. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

RF: Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. SC: Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. GD: Writing – review & editing, Supervision. MC: Writing – review & editing, Supervision. FP: Writing – review & editing, Supervision. IG: Writing – review & editing, Supervision, Methodology, Investigation, Data curation, Conceptualization. FB: Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We thank the study participants for volunteering their time and all the parents who gave their consent approval to the research.

Conflict of interest

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Abell, F., Happe, F., and Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cogn. Dev.* 15, 1–16. doi: 10.1016/S0885-2014(00)00014-9

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders: DSM-5 (Vol. 5, No. 5). Washington, DC: American Psychiatric Association.

Andreou, M., and Skrimpa, V. (2020). Theory of mind deficits and neurophysiological operations in autism spectrum disorders: a review. *Brain Sci.* 10:393. doi: 10.3390/brainsci10060393

Angeleri, R., Gabbatore, I., Bosco, F. M., Sacco, K., and Colle, L. (2016). Pragmatic abilities in children and adolescents with autism spectrum disorder: a study with the ABaCo battery. *Minerva Psich.* 57, 93–103.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/s1364-6613(00)01538-2

Baillargeon, R., Scott, R. M., and He, Z. (2010). False-belief understanding in infants. *Trends Cogn. Sci.* 14, 110–118. doi: 10.1016/j.tics.2009.12.006

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The "Reading the mind in the eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry Allied Discip.* 42, 241–251. doi: 10.1111/1469-7610.00715

Begeer, S., Malle, B. F., Nieuwland, M. S., and Keysar, B. (2010). Using theory of mind to represent and take part in social interactions: comparing individuals with high-functioning autism and typically developing controls. *Eur. J. Dev. Psychol.* 7, 104–122. doi: 10.1080/17405620903024263

Bosco, F. M., Colle, L., De Fazio, S., Bono, A., Ruberti, S., and Tirassa, M. (2009a). Thomas: an exploratory assessment of theory of mind in schizophrenic subjects. *Conscious. Cogn.* 18, 306–319. doi: 10.1016/j.concog.2008.06.006

Bosco, F. M., Colle, L., Pecorara, R. S., and Tirassa, M. (2006). ThOMAS, Theory of Mind Assessment Scale: uno strumento per la valutazione della teoria della mente. *Sist. intell.* 18, 215–242. doi: 10.1422/22502

Bosco, F. M., Colle, L., and Tirassa, M. (2009b). The complexity of theory of mind. *Conscious. Cogn.* 18, 323–324. doi: 10.1016/j.concog.2008.12.007

Bosco, F. M., Gabbatore, I., Tirassa, M., and Testa, S. (2016). Psychometric properties of the theory of mind assessment scale in a sample of adolescents and adults. *Front. Psychol.* 7:566. doi: 10.3389/fpsyg.2016.00566

Bosco, F. M., Gabbatore, I., and Tirassa, M. (2014). A broad assessment of theory of mind in adolescence: The complexity of mindreading. *Conscious Cogn.* 24, 84–97. doi: 10.1016/j.concog.2014.01.003

Bowler, D. M. (1992). "Theory of mind" in Asperger's syndrome. *J. Child Psychol. Psychiatry* 33, 877–893. doi: 10.1111/j.1469-7610.1992.tb01962.x

Brewer, N., Young, R. L., and Barnett, E. (2017). Measuring theory of mind in adults with autism spectrum disorder. *J. Autism Dev. Disord.* 47, 1927–1941. doi: 10.1007/s10803-017-3080-x

Colle, L., Gabbatore, I., Riberi, E., Borroz, E., Bosco, F. M., and Keller, R. (2019). Mindreading abilities and borderline personality disorder: a comprehensive assessment using the theory of mind assessment scale. *Psychiatry Res.* 272, 609–617. doi: 10.1016/j.psychres.2018.12.102

Congiu, S., Fadda, R., Doneddu, G., and Striano, T. (2016). Impaired representational gaze following in children with autism spectrum disorder. *Res. Dev. Disabil.* 57, 11–17. doi: 10.1016/j.ridd.2016.06.008

Deschrijver, E., and Palmer, C. (2020). Reframing social cognition: Relational versus representational mentalizing. *Psychol. Bull.* 146, 941–969. doi: 10.1037/bul0000302

Devine, R. T., and Lecce, S. (2021). "Why study theory of mind in middle childhood and adolescence?" in Theory of mind in middle childhood and adolescence: Integrating multiple perspectives. eds. R. T. Devine and S. Lecce (London: Routledge), 55–76.

Frith, U. (1994). Autism and theory of mind in everyday life. *Soc. Dev.* 3, 108–124. doi: 10.1111/j.1467-9507.1994.tb00031.x

Frith, U., and De Vignemont, F. (2005). Egocentrism, allocentrism, and Asperger syndrome. *Conscious. Cogn.* 14, 719–738. doi: 10.1016/j.concog.2005.04.006

Frith, U., and Happé, F. (1994). Autism: beyond "theory of mind". *Cognition* 50, 115–132. doi: 10.1016/0010-0277(94)90024-8

Gao, S., Wang, X., and Su, Y. (2023). Examining whether adults with autism spectrum disorder encounter multiple problems in theory of mind: a study based on meta-analysis. *Psychon. Bull. Rev.* 30, 1740–1758. doi: 10.3758/s13423-023-02280-8

Garfinkel, S. N., Tiley, C., O'Keeffe, S., Harrison, N. A., Seth, A. K., and Critchley, H. D. (2016). Discrepancies between dimensions of interoception in autism: implications for emotion and anxiety. *Biol. Psychol.* 114, 117–126. doi: 10.1016/j.biopsycho.2015.12.003

Happé, F. G. (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

Hill, E. L. (2004). Executive dysfunction in autism. *Trends Cogn. Sci.* 8, 26–32. doi: 10.1016/j.tics.2003.11.003

Howlin, P., and Magiati, I. (2017). Autism spectrum disorder: outcomes in adulthood. *Curr. Opin. Psychiatry* 30, 69–76. doi: 10.1097/YCO.0000000000000308

Jones, C. R. G., Simonoff, E., Baird, G., Pickles, A., Marsden, A. J. S., Tregay, J., et al. (2018). The association between theory of mind, executive function, and the symptoms of autism spectrum disorder. *Autism Res.* 11, 95–109. doi: 10.1002/aur.1873

Laghi, F., Cotugno, A., Cecere, F., Sirolli, A., Palazzoni, D., and Bosco, F. M. (2014). An exploratory assessment of theory of mind and psychological impairment in patients with bulimia nervosa. *Br. J. Psychol.* 105, 509–523. doi: 10.1111/bjop.12054

Laghi, F., Terrinoni, A., Cerutti, R., Fantini, F., Galosi, S., Ferrara, M., et al. (2016). Theory of mind in non-suicidal self-injury (NSSI) adolescents. *Conscious. Cogn.* 43, 38–47. doi: 10.1016/j.concog.2016.05.004

Livingston, L. A., Colvert, E., the Social Relationships Study TeamBolton, P., and Happé, F. (2019). Good social skills despite poor theory of mind: exploring compensation in autism spectrum disorder. *J. Child Psychol. Psyc.* 60, 102–110. doi: 10.1111/jcpp.12886

Livingston, L. A., and Happé, F. (2017). Conceptualising compensation in neurodevelopmental disorders: Reflections from autism spectrum disorder. *Neurosci. Biobehav. Rev.* 80, 729–742. doi: 10.1016/j.neubiorev.2017.06.005

Lord, C., Rutter, M., DiLavore, P. C., and Risi, S. (2005). Autism diagnostic observation schedule, (ADOS). Firenze: Giunti Organizzazioni Speciali.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Luyster, R. J., Gotham, K., et al. (2013). Autism diagnostic observation schedule-second edition (ADOS-2). Firenze: Hogrefe Editore.

Loukusa, S., Gabbatore, I., Kotila, A. R., Dindar, K., Mäkinen, L., Leinonen, E., et al. (2023). Non-linguistic comprehension, social inference and empathizing skills in autistic young adults, young adults with autistic traits and control young adults: group differences and interrelatedness of skills. *Int. J. Lang. Comm. Dis.* 58, 1133–1147. doi: 10.1111/1460-6984.12848

Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *Eur. J. Neurosci.* 47, 497–514. doi: 10.1111/ejn.13720

Murray, K., Johnston, K., Cunnane, H., Kerr, C., Spain, D., Gillan, N., et al. (2017). A new test of advanced theory of mind: the "strange stories film task" captures social processing differences in adults with autism spectrum disorders. *Autism Res.* 10, 1120–1132. doi: 10.1002/aur.1744

Nichols, S., and Stich, S. (2004). Reading One's own mind: self-awareness and developmental psychology. *Can. J. Philos.* 30, 297–339. doi: 10.1080/00455091.2004.10717609

Ratto, A. B., and Mesibov, G. B. (2015). Autism spectrum disorders in adolescence and adulthood: long-term outcomes and relevant issues for treatment and research. *Sci. China Life Sci.* 58, 1010–1015. doi: 10.1007/s11427-012-4295-x

Senju, A., Southgate, V., White, S., and Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science* 325, 883–885. doi: 10.1126/science.1176170

Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Stat. Methods Med. Res.* 7, 301–317. doi: 10.1177/096228029800700306

Sommerville, J. A., Bernstein, D. M., and Meltzoff, A. N. (2013). Measuring beliefs in centimeters: private knowledge biases preschoolers' and adults' representation of others' beliefs. *Child Dev.* 84, 1846–1854. doi: 10.1111/cdev.12110

Tirassa, M., Bosco, F. M., and Colle, L. (2006). Rethinking the ontogeny of mindreading. *Conscious. Cogn.* 15, 197–217. doi: 10.1016/j.concog.2005.06.005

Wechsler, D. (2012). WISC-IV Wechsler intelligence scale for children Quarta Edizione [Manuale di Somministrazione e Scoring]. Firenze: Giunti OS.

Wechsler, D. (2013). Wechsler Adult Intelligence Scale, Quarta Edizione [Manuale di Somministrazione e Scoring]. Firenze: Giunti OS.

Wellman, H. M. (2014). Making minds: how theory of mind develops. Oxford: Oxford University Press.

Wellman, H. M., and Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Dev.* 75, 523–541. doi: 10.1111/j.1467-8624.2004.00691.x

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128. doi: 10.1016/0010-0277(83)90004-5

Yu, C. L., and Wellman, H. M. (2023). All humans have a 'theory of mind'. *J. Autism Dev. Disord.* 53, 2531–2534. doi: 10.1007/s10803-022-05584-1

frontiers | Frontiers in Psychology

TYPE Perspective
PUBLISHED 23 December 2024
DOI 10.3389/fpsyg.2024.1507298

OPEN ACCESS

EDITED BY
Simone G. Shamay-Tsoory,
University of Haifa, Israel

REVIEWED BY
Ilaria Castelli,
University of Bergamo, Italy

*CORRESPONDENCE
Ilaria Gabbatore
✉ ilaria.gabbatore@unito.it
Francesca M. Bosco
✉ francesca.bosco@unito.it

RECEIVED 07 October 2024
ACCEPTED 09 December 2024
PUBLISHED 23 December 2024

CITATION
Gabbatore I, Bosco FM and Tirassa M (2024) What are they all doing in that restaurant? Perspectives on the use of theory of mind. *Front. Psychol.* 15:1507298.
doi: 10.3389/fpsyg.2024.1507298

What are they all doing in that restaurant? Perspectives on the use of theory of mind

Ilaria Gabbatore[1]*, Francesca M. Bosco[2,3]* and Maurizio Tirassa[3,4]

[1]Department of Humanities, GIPSI Research Group, University of Turin, Turin, Italy, [2]Department of Psychology, GIPSI Research Group, University of Turin, Turin, Italy, [3]Neuroscience Institute of Turin, University of Turin, Turin, Italy, [4]Department of Psychology, University of Turin, Turin, Italy

If "theory of mind" is conceived as reasoning in a strict sense, then it can be said to be useful only at certain times; however, this leaves the rest of social cognition hardly comprehensible. If "theory of mind" is used instead to refer to a mentalist ontology and the consequent awareness that we ourselves and the others function on mental states, then we need new approaches that explain the flow of social experience. To illustrate these points, we outline the general conceptual framework that underlies most empirical studies of theory of mind and discuss their pros and cons; then, we discuss the Theory of Mind Assessment Scale, a tool developed to investigate the complexity of theory of mind, which adopts a different perspective and has been successfully tested on numerous populations.

KEYWORDS

theory of mind, social cognition, mental states, assessment tools, Th.o.m.a.s.

1 Introduction

Most research on theory of mind participates in a common framework. The overall goal is to map the development of such faculty in the infancy and the childhood. Legitimate questions could be: what is an agent to a child? What kinds of entities do children perceive as agents, how, and why? What kinds of mental states and reasoning do children attribute to an agent, and through what kinds of reasoning of their own? Around what ages and through what steps do these developments take place? To find answers to these and other similar questions, different types of cognitive challenges are presented to children of different ages, either as problems explicitly posed by the experimenters or embedded in manufactured world situations that hopefully appeal to their spontaneous curiosity or desires. The responses or behaviors collected (whichever is required in each setting) are expected to provide information about the children's naturally emerging social cognition. The ideal challenge is one which, by the very fact of being solved, proves beyond reasonable doubt the presence of the relevant form of psychological reasoning.

Of course this is a sensible strategy, for historical and conceptual reasons as well as for applicative ones. Studying the early development of a faculty, especially one that is so crucial in ontogeny and phylogeny, may help shed light on what the child's cognitive endowment is before culture and individual experience become too important (Gabbatore et al., 2023). This research thus bears on the debate about nativism and the nature/nurture relationships and its more recent incarnation, namely the one about initial knowledge (e.g., Baillargeon et al., 2016). To understand the acquisition of theory of mind may also have important clinical correlates (especially for autism, which has been characterized as involving an impairment of theory of mind: Baron-Cohen, 1995; Happé and Frith, 1996) and pedagogical and educational applications (e.g., Grover, 2015; Lecce and Devine, 2022; Smogorzewska et al., 2020; Wang, 2015).

Because of this complex backdrop the goal, on the epistemological level and therefore on the methodological one, is to achieve the greatest possible clarity about what is going on in the child's mind. This is also true of the study of theory of mind in animals, which was the actual field of Premack and Woodruff's (1978) foundational article. The problem of how to distinguish between a "true" theory of mind and a "mere" expectation about another agent's behavior was immediately raised in Dennett's (1978) commentary on Premack and Woodruff's paper. Dennett argued that certainty about someone's capability of psychological reasoning can only be achieved if she can hold a negative belief about another agent's knowledge, not a positive one. In other words, if Tommy and I have identical knowledge of a certain state of affairs it will not be clear to an observer whether I interpret and predict Tommy's actions on the basis of his knowledge or mine; if, however, there is a knowable difference between Tommy's knowledge and mine it will be possible to draw such distinction, thus proving whether I am aware that he has mental states of his own which need not be identical to mine.

This line of reasoning also provided the basis for the famous papers by Baron-Cohen et al. (1985) about autism, whose title of course echoed that of Premack and Woodruff, and by Wimmer and Perner (1983), who devised the first false belief task to be employed with young children.

The rest, as they say, is history: the false belief task has had its ups and downs, other experimental paradigms have been devised, theories have been proposed and refined (see, e.g., Kulke et al., 2019; Onishi and Baillargeon, 2005), but the general research framework has not changed much; nor, given its apparent overall reasonableness, have there been particularly compelling reasons to change it. We do not have the space to discuss this rich area here; excellent reviews and systematizations have been published by Barone et al. (2019), Matthews et al. (2018), Poulin-Dubois et al. (2023), Schneider et al. (2017), and Wellman et al. (2001).

2 Beyond childhood

The framework we have outlined has proven precious both in developmental and in clinical psychology. It is useful in general for locating specific turning points in the development or the decay of theory of mind; however, unlike what happens in geometry, such points do not allow to extrapolate a curve, nor do they tell much about the actual nature and functioning of theory of mind. When a child passes a certain task, all we know is that she possesses (and uses) the ability to do so. This is clearly important and interesting, but does not exhaust the questions: what does the child do when she is not handling false beliefs? What actually is her theory of mind and how does she use it in her everyday life? What becomes of her theory of mind as she grows to be an adolescent, an adult, and an elderly person? What are the workings of social cognition in the human species? The capability of passing an experimental task does not smoothly translate into the capability of interacting in real-life social situations, both because each kind of activity embodies different cognitive demands and because of the roles that are possibly played by motivation, social status, and other contextual factors (Astington, 2003; Massaro and Castelli, 2009).

Furthermore, except in particular clinical contexts, the tasks and experiments suitable for young children lose much of their usefulness at different ages. Most of them, if proposed to an elder child or an adult, would have him think that the experimenter was making fun of him or that there was some hidden trick. Even a serious response would not be informative anyway: once someone has started passing a task, he will probably just continue to pass it for the rest of his life.

Subtler, more naturalistic tasks have therefore been designed for the study of the adolescent and the adult theory of mind. Some, like the Reading the Mind in The Eyes (Baron-Cohen et al., 2001), assess specific abilities and thus maintain the "punctiform" approach that characterizes children tasks; others, like the Faux Pas (Stone et al., 1998) or the Strange Stories (Happé, 1994) tasks, explore the participant's ability to handle the mental states surrounding some social mistake or blunder made by the fictional protagonists of short narratives. All focus on how the participants make sense of a scene of which they are spectators; while the specifics change from instance to instance, the underlying idea remains the same.

Other tools have been developed that focus on specific tasks or rely on video instead of narrative material, e.g., the Theory of Mind Picture Stories Task (Brüne, 2003), the Conversations and Intimations task (Ouellet et al., 2010), and the Virtual Assessment of Mentalising Ability (Canty et al., 2017). Karmakar and Dogra (2019) offer a review of the several tools available.

The Theory of Mind Assessment Scale (Th.o.m.a.s.; Bosco et al., 2016; Bosco et al., 2009a; Bosco et al., 2006), to whose development and application we and other colleagues have collaborated, takes a different stance. It consists in a semi-structured interview composed of 35–40 questions, lasting possibly around an hour. The number of questions is not rigid because if the interviewee spontaneously extends an answer to the contents of another question, the latter may be omitted. On the theoretical level, this tool views theory of mind as a complex, sophisticated faculty that humans employ for comprehending both a partner's mental states and those of their own and for planning an attempt to modify them (see also Bosco et al., 2009b). The capability of affecting the interlocutors' mental states is the foundation of human interactions (Tirassa, 1999; Tirassa and Bosco, 2008) and therefore requires first the capability of understanding what such states are, how they function, how they causally relate to each other and to the world, and what may affect them. It also requires to be able to distinguish the nature and functional role of at least a few basic mental states like beliefs, desires, intentions, or emotions. The interview explores all these aspects of theory of mind, gently pushing the interviewee to make explicit her awareness of the various issues involved. The transcript of the interview is assessed separately by two trained, independent judges on an established set of criteria; once any differences in assessment have been resolved, the final report provides a complex profile of the interviewee's theory of mind. The Th.o.m.a.s. thus embodies a theory of theory of mind which goes beyond the punctiform measurement of a single one of its component abilities.

Initially developed in Italian (Bosco et al., 2006), the Th.o.m.a.s. was translated into English (Bosco et al., 2016), validated (Bosco et al., 2016) and successfully employed with populations such as typically developing (Bosco et al., 2014b) and self-injury adolescents (Laghi et al., 2016), sex offenders (Castellino et al., 2011), young women with bulimia nervosa (Laghi et al., 2014), persons with schizophrenia (Bosco et al., 2009a), with congenital heart disease (Chiavarino et al., 2015), alcohol use disorder (Bosco et al., 2014a), border personality disorder (BPD; Colle et al., 2019), opiate dependency (Gandolphe et al., 2018), persons receiving treatment for non-psychotic disorders (Francesconi et al., 2016), and persons with medication-overuse headache and migraine (Romozzi et al., 2022). In such populations the Th.o.m.a.s. has allowed to highlight profiles of theory of mind impairment (e.g., Bosco et al., 2024). For example, individuals with bulimia nervosa (Laghi et al., 2014) found it harder to accurately answer Th.o.m.a.s. questions that asked to reflect on other persons' mental state (i.e., third-person ToM). In contrast, they found it easier to reason about their own mental states (i.e., first-person ToM). A similar pattern was identified in individuals with alcohol use disorder (Bosco et al., 2014a). In a related vein, people with borderline personality disorder (Colle et al., 2019) exhibited difficulties in Th.o.m.a.s. scales that evaluate the ability to attribute mental states from an allocentric perspective, i.e., one that is independent of one's own standpoint. However, they performed similarly to controls on the scales based on the egocentric perspective. Interestingly, this discrepancy between allocentric and egocentric mindreading abilities was not observed in persons with a diagnosis of schizophrenia (Bosco et al., 2009a), who performed equally poorly as healthy controls on the Th.o.m.a.s. scales assessing these perspectives.

To avoid oversimplifying these results here, we refer interested readers to the specific papers for more details on the pattern of strengths and weaknesses across the various populations mentioned.

3 Humans as full-time mentalists

The Th.o.m.a.s. investigates the interviewee's retrospective awareness of her theory of mind as it is generated at the time of the interview and stimulated by the interview itself. A brief discussion may be necessary.

Consider: as I chat idly with an old friend in a pub, do I need to engage in any reasoning to understand what she is saying? There is an obvious sense in which I (mostly) do not; yet, it is equally obvious that I am not viewing her (and the other patrons, the staff, and myself) like a behaviorist would want me to (Skinner, 1938; Watson, 1913), nor am I oscillating between behaviorism and mentalism or finding myself on some middle ground between the two. I just know my friend's character, the general lines and many details of her life, her way of thinking and so on, and I interact with her accordingly. My comprehension of what she says comes in fully psychological terms even though I am not specifically reasoning about her mental states, or even wondering what they might be. Yet, I can always ask myself, more or less intensely, how she really feels about a certain matter or what the intents could be of a common acquaintance she is telling me about. In doing or not doing so, or doing so to a certain depth, I do not become more or less mentalist: I always am, but I dedicate variable amounts of time, attention and effort to actually reasoning about her thoughts, depending on how I sense the situation. This is only a matter of circumstances: it has nothing to do with having theory of mind or not or being a behaviorist; it has to do with the ways, the extent, the goals etc. in which I am using my knowledge of the mind.

This is just an instance of how consciousness and the mind always work. When I walk in the street I do not usually reason about colors; yet I see the world colored. I can pay cursory attention to the traffic lights: this is not really reasoning about them, it is just a slightly higher level of attention than I generally pay to the colors of the dress of the people I pass. I can engage in actual reasoning, e.g., when I start looking at the clothes in a shop, wondering what will go best with a certain dress I have at home. This may even become difficult, e.g., if there is not much light and I try to realize what color a certain dress really is. Yet, nobody would suggest that, when I am not reasoning about colors, I only see in black and white or in shades of grey.

The same applies to action. The extent to which our movements or speech are conscious and deliberate depends on what we are doing and why. When it is my turn to tell my friend what I have been up to since we last saw each other, I will probably not painstakingly choose each and every word to pronounce: I will just follow the thread of the conversation, taking care of the general sense of it and counting on her to understand it. However, if the topic shifts to something that I know troubles her, suddenly I will become much more careful about the possible effects of my words.

Any number of examples could be made. When I pass along the window of a restaurant, do I (normally) reason about what might have pushed all those people to get in or will I just take it for granted that they are hungry? Yet hunger is undoubtedly a mental state. Once again, my theory of mind is just revved down, so to speak, but never turned off, and always ready to return to full operation.

This is obvious in everyday life, but hard to capture in theoretical terms. Yet, we believe this point is crucial for the cognitive sciences in general and for this debate in particular. Not all that is mental is reasoning; not all that is mental is problem solving.

In retrospect, however, we are typically capable of summarizing sequences of events or thoughts and giving an average assessment of the quality and depth of our own and others' performance as well as of the underlying cognitive framework and the results achieved.

4 Conclusion

To test a subject's capability of handling one or more well-structured theory of mind problems does certainly say something. It is the only way we can achieve certainty of the presence of theory of mind in the cognitive architecture of other species. In ours, it may be useful for diagnostic purposes, if a theory exists that appropriately links those measures to the condition investigated, a bit like it is done with glycaemia and diabetes. For the same reasons and under the same conditions, it may help map the faculty's development during childhood or its decay under specific conditions.

However, this strategy may be less informative about other issues. It has nothing to say about the social cognition of an agent who does not pass the relevant tasks, and is generally unable to provide a wider description of the workings of theory of mind. Some persons might pass all the tasks and still fail to decipher their spouse's thoughts or to realize that they are being deceived in everyday circumstances; conversely, a child might get by just fine in everyday social life but struggle to solve the abstract problems. To equate theory of mind with the ability to pass certain chosen tasks also compels to lower as much as possible the age at which children become able to do so, both because that is the only description available of their social competence and because there is an implication that a child who does not solve the tasks can only be a behaviorist, which there are several reasons not to accept (Bosco and Tirassa, 1998). This is true of young children in general (Tirassa et al., 2006a) and in particular of the autistic ones, whose differences to typically developing children and subsequent possible development of a functioning social life become essentially incomprehensible.

This strategy may also prevent the exploration of other possibilities both for infant cognition (based, e.g., on intersubjectivity: Airenti, 2015; Trevarthen, 1998, or on a basic notion of sharedness: Tirassa et al., 2006a, 2006b) and for adolescence (e.g., Brizio et al., 2015). On a more contingent level, many existing tasks appear to favor an individualistic and spectatorial approach over one of sharing and participation and to limit the definition of theory of mind to the comprehension of an observed problem. However, theory of mind is much more than this, and even comprehension is more radically based in interaction than in mere observation (Trevarthen, 1998).

Therefore, it may be desirable to explore how people construe, describe and criticize their own theory of mind and that of the others, how they use it to capture and understand the causal relationships between mental states and between mental states and the world, and how they practically employ such knowledge to achieve actual changes in a given situation. This is what the Th.o.m.a.s. does. On the other hand, it requires the interviewee to be capable of sustaining the interview and to possess at least a working level of social awareness and expertise. Thus, until tools are developed that capture the best of the two worlds, a trade-off appears to exist between different approaches to the matter.

Since it functions in retrospect, leveraging on the interviewees' recapitulation of their past experience, the Th.o.m.a.s. also avoids reducing social life to a sequence of formal problems to observe, reason about, and solve, interspersed with intervals that either remain incomprehensible or can only be described as behavioristic. Humans experience their social life as a continuous flow of thoughts, actions and events, always based on a mentalist ontology, whose workings include occasional bouts of actual reasoning when needed, with variable degrees of commitment and difficulty (and variable success). This special issue asks *when and how theory of mind is useful*. We believe that the answer depends less on how human beings function than on how theory of mind is defined. If theory of mind is conceived as reasoning in a strict sense, then it can be said to be useful only at certain times; this, however, leaves the rest of social life hardly comprehensible. However, if theory of mind is used to refer to a continuous mentalist ontology and the consequent awareness that we ourselves and the others function on mental states, then we need new approaches to study, describe and explain the flow of social experience and the ways in which we treat the problems that occasionally surface from it. In the meantime, there exists at least one instrument that allows a thorough exploration of the matter.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

IG: Writing – original draft. FB: Conceptualization, Funding acquisition, Writing – original draft. MT: Conceptualization, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was supported by PRIN 2022, Prot. n. 2022CZF8KA, project title: "ACTIVe communication: Assessment and enhancement of pragmatic and narrative skills in hEaLthY aging (ACTIVELY)" Avviso pubblico n. 104 del 02/02/2022 - PRIN 2022 PNRR M4C2 Inv. 1.1. Ministero dell'Università e della Ricerca (Financed by EU, NextGenerationEU) - CUP G53D23003110006.

Conflict of interest

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

References

Airenti, G. (2015). Theory of mind: a new perspective on the puzzle of belief ascription. *Front. Psychol.* 6:1184. doi: 10.3389/fpsyg.2015.01184

Astington, J. W. (2003). "Sometimes necessary, never sufficient: false belief understanding and social competence" in Individual differences in theory of mind: Implications for typical and atypical development. eds. B. Repacholi and V. Slaughter (New York, NY: Psychology Press), 13–38.

Baillargeon, R., Scott, R. M., and Bian, L. (2016). Psychological reasoning in infancy. *Annu. Rev. Psychol.* 67, 159–186. doi: 10.1146/annurev-psych-010213-115033

Baron-Cohen, S. (1995). Mindblindness. Cambridge, MA: MIT Press.

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition* 21, 37–46. doi: 10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The "'Reading the mind in the eyes'" test revised version: a study with Normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry Allied Discip.* 42, 241–251. doi: 10.1111/1469-7610.00715

Barone, P., Corradi, G., and Gomila, A. (2019). Infants' performance in spontaneous-response false belief tasks: a review and meta-analysis. *Infant Behav. Dev.* 57:101350. doi: 10.1016/j.infbeh.2019.101350

Bosco, F. M., Capozzi, F., Colle, L., Marostica, P., and Tirassa, M. (2014a). Theory of mind deficit in subjects with alcohol use disorder: an analysis of mindreading processes. *Alcohol Alcohol.* 49, 299–307. doi: 10.1093/alcalc/agt148

Bosco, F. M., Colle, L., De Fazio, S., Bono, A., Ruberti, S., and Tirassa, M. (2009a). Th.O.M.A.S.: an exploratory assessment of theory of mind in schizophrenic subjects. *Conscious. Cogn.* 18, 306–319. doi: 10.1016/j.concog.2008.06.006

Bosco, F. M., Colle, L., Pecorara, R. S., and Tirassa, M. (2006). ThOMAS, Theory of Mind Assessment Scale: uno strumento per la valutazione della teoria della mente. *Sistemi Intelligenti* XVIII, 215–242. doi: 10.1422/22502

Bosco, F. M., Colle, L., Salvini, R., and Gabbatore, I. (2024). A machine-learning approach to investigating the complexity of theory of mind in individuals with schizophrenia. *Heliyon* 10:e30693. doi: 10.1016/j.heliyon.2024.e30693

Bosco, F. M., Colle, L., and Tirassa, M. (2009b). The complexity of theory of mind. *Conscious. Cognit.* 18, 323–324. doi: 10.1016/j.concog.2008.12.007

Bosco, F. M., Gabbatore, I., and Tirassa, M. (2014b). A broad assessment of theory of mind in adolescence: the complexity of mindreading. *Conscious. Cogn.* 24, 84–97. doi: 10.1016/J.CONCOG.2014.01.003

Bosco, F. M., Gabbatore, I., Tirassa, M., and Testa, S. (2016). Psychometric properties of the theory of mind assessment scale in a sample of adolescents and adults. *Front. Psychol.* 7:566. doi: 10.3389/FPSYG.2016.00566/BIBTEX

Bosco, F. M., and Tirassa, M. (1998). "Sharedness as an innate basis for communication in the infant" in Proceedings of the 20th annual conference of the cognitive science society. eds. M. A. Gernsbacher and S. J. Derry (Mahwah, NJ: Erlbaum), 162–166.

Brizio, A., Gabbatore, I., Tirassa, M., and Bosco, F. M. (2015). "No more a child, not yet an adult": studying social cognition in adolescence. *Front. Psychol.* 6:1011. doi: 10.3389/fpsyg.2015.01011

Brüne, M. (2003). Theory of mind and the role of IQ in chronic disorganized schizophrenia. *Schizophr. Res.* 60, 57–64. doi: 10.1016/S0920-9964(02)00162-7

Canty, A. L., Neumann, D. L., Fleming, J., and Shum, D. H. (2017). Evaluation of a newly developed measure of theory of mind: the virtual assessment of mentalising ability. *Neuropsychol. Rehab.* 27, 834–870. doi: 10.1080/09602011.2015.1052820

Castellino, N., Bosco, F. M., Marshall, W. L., Marshall, L. E., and Veglia, F. (2011). Mindreading abilities in sexual offenders: an analysis of theory of mind processes. *Conscious. Cogn.* 20, 1612–1624. doi: 10.1016/j.concog.2011.08.011

Chiavarino, C., Bianchino, C., Brach-Prever, S., Riggi, C., Palumbo, L., Bara, B. G., et al. (2015). Theory of mind deficit in adult patients with congenital heart disease. *J. Health Psychol.* 20, 1253–1262. doi: 10.1177/1359105313510337

Colle, L., Gabbatore, I., Riberi, E., Borroz, E., Bosco, F. M., and Keller, R. (2019). Mindreading abilities and borderline personality disorder: a comprehensive assessment using the theory of mind assessment scale. *Psychiatry Res.* 272, 609–617. doi: 10.1016/j.psychres.2018.12.102

Dennett, D. C. (1978). Beliefs about beliefs. *Behav. Brain Sci.* 1, 568–570. doi: 10.1017/S0140525X00076664

Francesconi, M., Minichino, A., Carrión, R. E., Chiaie, R. D., Bevilacqua, A., Parisi, M., et al. (2016). Theory of mind as a mediator variable between neurocognition and functioning in young individuals in treatment with secondary services for non-psychotic disorders. *Psychiatry Res.* 246, 415–420. doi: 10.1016/J.PSYCHRES.2016.09.057

Gabbatore, I., Dindar, K., Pirinen, V., Vähänikkilä, H., Mämmelä, L., Kotila, A., et al. (2023). Silent Finns and talkative Italians? An investigation of communicative differences and similarities as perceived by parents in typically developing children. *First Lang.* 43, 313–335. doi: 10.1177/01427237221149310

Gandolphe, M. C., Lecluyse, B., Triquet, C., Brunelle, E., Duparcq, J. P., and Nandrino, J. L. (2018). Mind reading abilities in opiate-dependent patients: an exploratory study. *Compr. Psychiatry* 83, 46–52. doi: 10.1016/J.COMPPSYCH.2018.03.001

Grover, V. (2015). Theory of mind: concept and application for classroom learning. *Eur. Acad. Res.* II, 13050–13060.

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

Happé, F., and Frith, U. (1996). The neuropsychology of autism. *Brain* 119, 1377–1400. doi: 10.1093/BRAIN/119.4.1377

Karmakar, A., and Dogra, A. K. (2019). Assessment of theory of mind in adults: beyond false belief tasks. *Act Nerv Super* 61, 142–146. doi: 10.1007/S41470-019-00028-1

Kulke, L., Wübker, M., and Rakoczy, H. (2019). Is implicit theory of mind real but hard to detect? Testing adults with different stimulus materials. *R. Soc. Open Sci.* 6:190068. doi: 10.1098/rsos.190068

Laghi, F., Cotugno, A., Cecere, F., Sirolli, A., Palazzoni, D., and Bosco, F. M. (2014). An exploratory assessment of theory of mind and psychological impairment in patients with bulimia nervosa. *Br. J. Psychol.* 105, 509–523. doi: 10.1111/bjop.12054

Laghi, F., Terrinoni, A., Cerutti, R., Fantini, F., Galosi, S., Ferrara, M., et al. (2016). Theory of mind in non-suicidal self-injury (NSSI) adolescents. *Conscious. Cogn.* 43, 38–47. doi: 10.1016/J.CONCOG.2016.05.004

Lecce, S., and Devine, R. T. (2022). Theory of mind at school: academic outcomes and the influence of the school context. *Infant Child Dev.* 31:e2274. doi: 10.1002/ICD.2274

Massaro, D., and Castelli, I. (2009). Mentalization in communicative and socio-relational interactions: considerations about a theory-of-mind modelling. *Stu. Commun. Sci.* 9, 103–130.

Matthews, D., Biney, H., and Abbot-Smith, K. (2018). Individual differences in Children's pragmatic ability: a review of associations with formal language, social cognition, and executive functions. *Lang. Learn. Dev.* 14, 186–223. doi: 10.1080/15475441.2018.1455584

Onishi, K. H., and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* 1979, 255–258. doi: 10.1126/SCIENCE.1107621

Ouellet, J., Scherzer, P. B., Rouleau, I., Métras, P., Bertrand-gauvin, C., Djerroud, N., et al. (2010). Assessment of social cognition in patients with multiple sclerosis. *J. Int. Neuropsychol. Soc.* 16, 287–296. doi: 10.1017/S1355617709991329

Poulin-Dubois, D., Goldman, E. J., Meltzer, A., and Psaradellis, E. (2023). Discontinuity from implicit to explicit theory of mind from infancy to preschool age. *Cogn. Dev.* 65:101273. doi: 10.1016/J.COGDEV.2022.101273

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/S0140525X00076512

Romozzi, M., Di Tella, S., Rollo, E., Quintieri, P., Silveri, M. C., Vollono, C., et al. (2022). Theory of mind in migraine and medication-overuse headache: a cross-sectional study. *Front. Neurol.* 13:968111. doi: 10.3389/fneur.2022.968111

Schneider, D., Slaughter, V. P., and Dux, P. E. (2017). Current evidence for automatic theory of mind processing in adults. *Cognition* 162, 27–31. doi: 10.1016/j.cognition.2017.01.018

Skinner, B. (1938). The behavior of organisms: An experimental analysis. New York, NY: Appleton-Century.

Smogorzewska, J., Szumski, G., and Grygiel, P. (2020). Theory of mind goes to school: does educational environment influence the development of theory of mind in middle childhood? *PLoS One* 15:e0237524. doi: 10.1371/journal.pone.0237524

Stone, V. E., Baron-Cohen, S., and Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *J. Cogn. Neurosci.* 10, 640–656. doi: 10.1162/089892998562942

Tirassa, M. (1999). Communicative competence and the architecture of the mind/brain. *Brain Lang.* 68, 419–441. doi: 10.1006/BRLN.1999.2121

Tirassa, M., and Bosco, F. M. (2008). "On the nature and role of Intersubjectivity in human communication" in Enacting Intersubjectivity: A cognitive and social perspective on the study of interactions. eds. F. Morganti, A. Carassa and G. Rova (Amsterdam, NE: IOS Press), 81–95.

Tirassa, M., Bosco, F. M., and Colle, L. (2006a). Rethinking the ontogeny of mindreading. *Conscious. Cogn.* 15, 197–217. doi: 10.1016/J.CONCOG.2005.06.005

Tirassa, M., Bosco, F. M., and Colle, L. (2006b). Sharedness and privateness in human early social life. *Cogn. Syst. Res.* 7, 128–139. doi: 10.1016/J.COGSYS.2006.01.002

Trevarthen, C. (1998). "The concept and foundations of infant intersubjectivity" in Intersubjective communication and emotion in early ontogeny. ed. S. Bråten (Cambridge: Cambridge University Press), 15–46.

Wang, Z. (2015). Theory of mind and children's understanding of teaching and learning during early childhood. *Cogent Educ.* 2:1011973. doi: 10.1080/2331186X.2015.1011973

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychol. Rev.* 20, 158–177. doi: 10.1037/h0074428

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

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128. doi: 10.1016/0010-0277(83)90004-5

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 07 February 2025
DOI 10.3389/fpsyg.2025.1535722

Check for updates

OPEN ACCESS

EDITED BY
Sara Isernia,
Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy

REVIEWED BY
Valeria Manera,
Université Côte d'Azur, France
Valeria Blasi,
Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy

*CORRESPONDENCE
Claudia Casadio
✉ claudia.casadio@unimore.it

†These authors share first authorship

RECEIVED 27 November 2024
ACCEPTED 20 January 2025
PUBLISHED 07 February 2025

CITATION
Sola C, Zanelli V, Molinari MA, Casadio C, Ricci F, Carpentiero O, Tondelli M, Lui F, Nichelli PF and Benuzzi F (2025) Understanding basic and social emotions in Alzheimer's disease and frontotemporal dementia. *Front. Psychol.* 16:1535722. doi: 10.3389/fpsyg.2025.1535722

Understanding basic and social emotions in Alzheimer's disease and frontotemporal dementia

Carlotta Sola[1†], Vanessa Zanelli[2†], Maria Angela Molinari[3], Claudia Casadio[2*], Francesco Ricci[2], Omar Carpentiero[2], Manuela Tondelli[2], Fausta Lui[2], Paolo Frigio Nichelli[2] and Francesca Benuzzi[2]

[1]Physical Medicine and Rehabilitation Unit, AUSL-IRCCS di Reggio Emilia, Reggio Emilia, Italy, [2]Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy, [3]Azienda Ospedaliero Universitaria (AOU) of Modena, Modena, Italy

Introduction: Recent developments in the field of social cognition have led to a renewed interest in basic and social emotion recognition in early stages of Alzheimer's Disease (AD) and FrontoTemporal Dementia (FTD). Despite the growing attention to this issue, only few studies have attempted to investigate emotion recognition using both visual and vocal stimuli. In addition, recent studies have presented conflicting findings regarding the extent of impairment in patients in the early stages of these diseases. The present study aims to investigate emotion understanding (both basic and social emotions), using different tasks with visual and auditory stimuli, to identify supramodal deficits in AD and FTD to provide a reliable tool to better outline their behavioral and emotional profile and useful instruments for their management.

Methods: Eighteen patients with AD and 15 patients with FTD were included in the study. Healthy control (HCs) subjects were recruited to obtain normative data for basic emotion recognition tests and social emotion recognition tasks. To evaluate basic emotion recognition, the Facial Emotion Recognition Battery (FERB) and the Emotional Prosody Recognition Battery (EPRB) were administered. To evaluate social emotion recognition, the Faux Pas (FP), Reading the Mind in the Eyes (RME), and Reading the Mind in the Voice (RMV) tests were employed.

Results: FTD patients performed significantly worse than HCs in most of the subtests of the basic emotion recognition batteries, where, instead, AD patients were significantly impaired only when required to match emotional facial expression in different individuals (subtask of the FERB). Moreover, FTD patients scored significantly lower in RME and RMV tests compared both to AD patients and to HCs. In addition, ADs were selectively impaired in RMV as respect to HCs.

Discussion: FTD patients showed deficits in emotion recognition, affecting both basic and social emotions, whether conveyed through facial expressions or prosody. This result may explain the well-known social behavioral difficulties observed in FTD patients from the early stages of the disease. The fewer and specific deficits in AD patients with comparable MMSE scores may be attributed to the mild degree of impairment, as these deficits may appear later in the progression of AD.

KEYWORDS

Theory of Mind, emotion recognition, emotional prosody, Alzheimer's disease, frontotemporal dementia

1 Introduction

The term "Social Cognition" refers to several abilities involved in social information processing, consisting of inferring emotions and socially relevant stimuli to modulate behavior (Adolphs, 1999; Frith, 2008).

Emotional processing plays an important role among high-level social abilities; several studies support the idea that it relies on a broad neural network including fusiform face area (FFA; Kanwisher and Yovel, 2006), amygdala (Adolphs et al., 1998, 1994; Todorov et al., 2013), insula (Wicker et al., 2003; Craig, 2009). Additional areas appear to be specifically involved in prosody, an important social signal, during emotional recognition, including superior temporal sulcus (STS; Grandjean et al., 2005; Sander et al., 2005), middle temporal gyrus (MTG), inferior frontal gyrus (IFG; Johnstone et al., 2006), putamen, pallidum, subthalamic nucleus, and cerebellum (Ceravolo et al., 2021).

Recent data have emphasized the need for a supramodal approach to understanding the neural basis of emotion processing (Schirmer and Adolphs, 2017). Each distinct input channel engages partly non-overlapping neuroanatomical systems with different processing specializations. Then, elaborations of signals across different modalities converge into supramodal representations in areas involving a modality-non-specific abstract code, such as STS, prefrontal and posterior cingulate cortex.

Deficits in emotional processing are observed in both Alzheimer's Disease (AD) and Frontotemporal Dementia (FTD). In AD, mild impairments in emotion recognition, particularly for low-intensity or negative emotions, emerge early and worsen over time (Luzzi et al., 2007; Maki et al., 2013; Torres et al., 2015; Garcia-Cordero et al., 2021; Amlerova et al., 2022; Chaudhary et al., 2022). These deficits extend to multiple sensory modalities, including emotional prosody, likely due to overlap between memory and emotional processing regions affected by neurodegeneration (Bediou et al., 2012). On the other hand, in FTD, significant impairments in recognizing visual and vocal emotional stimuli, especially negative emotions, are more severe than in AD (Fernandez-Duque and Black, 2005; Dara et al., 2013; Bertoux et al., 2015; Bora et al., 2016; Jiskoot et al., 2021; Wright et al., 2018). These deficits, prominent in the behavioral variant (bvFTD), are linked to atrophy in brain regions involved in emotional and social cognition (Rascovsky et al., 2011).

Within social cognition abilities, Theory of Mind (ToM) pertains to the capacity to attribute mental states to others and to anticipate, describe, and elucidate behavior based on these mental states (Baron-Cohen, 1997). Traditionally, ToM is divided in two subcomponents: cognitive ToM and affective ToM (Zhou et al., 2023), which rely on different neural networks. The capacity to understand others' beliefs, intentions and goals (cognitive ToM; Amodio and Frith, 2006) has been connected to the activity of the dorsolateral prefrontal cortex (dlPFC; Shamay-Tsoory and Aharon-Peretz, 2007). On the other hand, the ventromedial prefrontal cortex (vmPFC) and the orbitofrontal cortex (OFC), together with the amygdala, are involved in the representation and top-down regulation of emotional states and represent the node for the affective processing of others' mental states (Abu-Akel and Shamay-Tsoory, 2011).

Whereas, cognitive ToM has been explored using the first-order (Baillargeon et al., 2010) and the second-order (Perner and Wimmer, 1985) false belief tasks, affective ToM has been usually investigated by using the Reading the Mind in the Eyes task (RME; Baron-Cohen et al., 2001) and the Reading the Mind in the Voice task (RMV; Rutherford et al., 2002; Golan et al., 2007). Additionally, the Faux Pas test (Baron-Cohen et al., 1999) is commonly used to assess ToM in a non-specific way.

Despite the extensive research on ToM abilities in neurodegenerative diseases, the findings are highly heterogeneous. Although some studies have shown that AD patients exhibit deficits only in ToM tasks that require high cognitive demand (Castelli et al., 2011; Demichelis et al., 2020; Kessels et al., 2021; de Lucena et al., 2023), other results suggest that certain subcomponents of ToM abilities are preserved in AD (e.g., interpretation of sarcasm, social inference, and emotion evaluation; Kumfor et al., 2017). In contrast, research on patients with FTD has shown more consistent results, with a widespread and severe impairment of ToM abilities, which could serve as a clinical marker distinguishing FTD from other neurodegenerative diseases (Gossink et al., 2018; Dodich et al., 2021).

Given the clinical and social importance of AD and FTD, and the relevance of social cognition in these two neurodegenerative diseases, the purpose of this study was to better characterize them, by using a complete assessment to investigate both visual and auditory processing, both for basic and for social emotions, in the same patients. Although these tests may not reveal such striking differences that can be used for individual diagnosis, we aim to provide a reliable tool to better outline the behavioral and emotional profile of these two pathologies, thus also providing useful instruments for their management. To this aim, we used both visual and prosodic stimuli, specifically, two batteries (the Facial Affect Recognition Battery and the Prosodic Affect Recognition Battery) devised by our research group (Benuzzi et al., 2004; Ariatti et al., 2008). Furthermore, processing of social emotions (affective ToM) within the visual and the prosodic domain was assessed by the RME and RMV tasks. Finally, the Faux Pas test (FP) was used to assess the cognitive component of ToM abilities.

We hypothesized that early stages FTD patients would exhibit a global impairment on emotion recognition tasks and in the affective component of ToM abilities, as opposed to substantially preserved functions in early stages AD patients.

2 Materials and methods

2.1 Participants

Patients affected by either AD or FTD were recruited among those followed by the Neuropsychology Service of the University Hospital (AOU) of Modena. Eighteen patients with AD (mean age = 72.8 years, SD ± 4.8 years; mean school age = 7.3 years, SD ± 4.5 years; mean MMSE = 25.4, SD ± 3.7) and 15 patients with FTD (mean age = 65.9 years, SD ± 8.7 years; mean school age = 8.4 years, SD ± 4.5 years; mean MMSE = 25.3, SD ± 4.9) were included in the study. They were all right-handed (assessed using the Edinburgh Inventory; Oldfield, 1971) and diagnosed

TABLE 1 Demographic and clinical data of the study population.

	Age (years)		Education level (years)		Gender		MMSE (corrected)	
	M	SD	*M*	SD	Male	Female	*M*	SD
AD	72.8	4.8	7.3	4.5	6	12	25.4	3.7
FTD	65.9	8.7	8.4	4.5	11	4	25.3	4.9
HC (emotion recognition battery)	68	7.7	8.5	3.9	23	46		
HC (ToM tasks)	67.6	12.7	9.4	4.6	7	13		

AD, Alzheimer's Disease; FTD, FrontoTemporal Dementia; HC, Healthy Controls; M, mean; SD, standard deviation; MMSE, Mini Mental State Examination.

with either AD or FTD according to criteria given by McKhann et al. (2011). Exclusion criteria were as follows: MMSE < 16, history of stroke, history of psychiatric illness or of traumatic brain injury.

As control groups, 70 healthy controls (HC; mean age = 68 years, SD ± 7.7 years; mean school age = 8.5 years, SD ± 3.9 years) were recruited through public announcements among employees or former employees of the University of Modena and Reggio Emilia. They were administered the emotion recognition battery. Among these HC participants, 20 were also submitted to the ToM tasks (mean age = 67.6 years, SD ± 12.7 years; mean school age = 9.4 years, SD ± 4.6 years). HCs were recruited according to the following exclusion criteria: no history of neurological or psychiatric diseases, alcoholism, brain injury, cerebrovascular disease or other neurological conditions. Moreover, exclusion criteria included the presence of depression and obsessive-compulsive disorders, since it has been demonstrated that these disorders interfere with emotion identification (Gur et al., 1992; Abbruzzese et al., 1995; Bouhuys et al., 1999). The presence of these diseases was assessed through the Beck Depression Inventory (BDI > 11; Sica and Ghisi, 2007) and a reduced version of Maudsley Obsessive-Compulsive Questionnaire (MOCQ-R < 75th percentile; Sanavio et al., 1986). See Table 1 for the demographic and clinical features of groups (AD, FTD, HC).

All subjects gave their informed consent to participate in the study. Consent was obtained according to the Declaration of Helsinki. Moreover, the study procedure was approved by the ethical committee of the University of Modena and Reggio Emilia (Comitato Etico di Ateneo per la Ricerca, CEAR; Prot. n 83243). Both patients and HC underwent a clinical neuropsychological evaluation which was conducted during a single session lasting ~1 h and a half. On a subsequent day, experimental tests (emotion recognition batteries and ToM tasks) were administered in a single session lasting between 40 min and 60 min, depending on the individual patients' abilities.

2.2 Materials

2.2.1 Basic emotion recognition

To evaluate the ability to process basic emotion (fear, happiness, sadness, anger, and disgust; Ekman, 1999), two batteries were used (Benuzzi et al., 2004).

The Facial Emotion Recognition Battery includes the following subtests:

Face Matching (FM). In this task, subjects are presented with a vertically arranged set of four neutral expression faces and must select the photograph identical to a target face. Photographs of different individuals of the same gender are used as distractors. The task includes 14 trials and assesses perceptual deficits in face discrimination, for each correct answer, one point was assigned (range score 0–14), thus the higher the score, the better the performance.

Facial Identity Recognition (FIR). This task evaluates the ability to recognize a single person across various facial expressions. It consists of 14 trials, in which the subject is asked to identify the target person from a vertically arranged set of four faces, each showing different expressions. The task assesses associative deficits in face perception. For each correct answer, one point was assigned (range score 0–14), thus the higher the score, the better the performance.

Facial Affect Naming (FAN). Subjects are asked to choose the name that best describes the emotional expression displayed from five options printed below a stimulus face. The subtest includes 25 trials, with five trials for each basic emotion, for each correct answer, one point was assigned (range score 0–25), thus the higher the score, the better the performance.

Facial Affect Selection (FAS). The participant is asked to select the face with an expression that matches a target label from a vertically arranged set of five ones. The test includes 25 trials, for each correct answer, one point was assigned (range score 0–25), thus the higher the score, the better the performance.

Facial Affect Matching (FAM). In this task, subjects must choose from a vertically arranged set of five faces, the one displaying the same expression as a stimulus face. The person in the stimulus photo is always different, with one identity foil included, i.e., a photograph of the same individual as the stimulus, but with a different expression. The test comprises 25 trials, for each correct answer, one point was assigned (range score 0–25), thus the higher the score, the better the performance.

FM and FIR subtests represent control tasks, since they assess the ability to discriminate the perceptual features of faces. On the other hand, FAN, FAS, and FAM assess basic emotion processing and recognition. There is no time limit to complete the task.

The Emotional Prosody Recognition Battery (Benuzzi et al., 2004; Ariatti et al., 2008; Bonora et al., 2011) evaluates the ability to process basic emotions from prosodic cues presented via a computer application. Before the administration of the battery, all participants underwent an auditory acuity evaluation. All subjects (HC, AD, and FTD) had a normal hearing threshold. Stimuli are brief Italian sentences with a neutral meaning (e.g., "Marta is combing the cat"). Sentences vary only with respect to emotional prosody, which could express one of the five basic emotions: fear, anger, sadness, happiness, and disgust. At the beginning of the testing session, the computer volume is regulated by the examiner according to the subject's requests. Sentences are presented both orally and in written form on a computer screen at the same time, and subjects can listen to each trial up to three times. The Emotional Prosody Recognition Battery includes different subtests as follows:

Vocal Identity Discrimination (VID) assesses basic voice discrimination abilities. Participants are asked to determine whether two sentences are spoken by the same person. VID consists of 16 pairs of neutral (aprosodic) stimuli, for each correct answer, one point was assigned (range score 0–16), thus the higher the score, the better the performance.

Prosodic Discrimination (PrD) measures basic intonation discrimination abilities. Given two sentences, subjects must identify whether they are uttered with the same prosodic intonation. PrD consists of 16 pairs of sentences expressing four different intonations: interrogative, declaratory, exclamatory, and imperative, for each correct answer, one point was assigned (range score 0–25), thus the higher the score, the better the performance.

Prosodic Affect Naming (PrAN) assesses emotional prosodic recognition abilities. Subjects are asked to choose from five options on the screen (representing five basic emotions) the one that best describes the emotional prosody of the target recorded sentence. PrAN consists of 25 trials, for each correct answer, one point was assigned (range score 0–25), thus the higher the score, the better the performance.

Prosodic Affect Discrimination (PrAD) measures emotional prosodic discrimination abilities. Given two recorded sentences, subjects must decide whether they are spoken with the same emotional prosody. PrAD consists of 45 pairs of sentences expressing the five basic emotions, for each correct answer, one point was assigned (range score 0–45), thus the higher the score, the better the performance.

Similarly to what happens for the Facial Emotion Recognition Battery, some sub-tests (here, VID, and PrD) represent control tasks, since they assess basic prosodic recognition abilities. On the contrary, PrAN and PrAD assess the emotional prosodic discrimination ability. There is no time limit for answering.

2.2.2 Social emotion recognition (ToM tasks)

To assess social emotion recognition, the following three tasks were used: the Reading the Mind in the Eyes test (RME; Baron-Cohen et al., 2001), the Reading the Mind in the Voice (RMV; Rutherford et al., 2002; Golan et al., 2007) and the Faux Pas test (FP; Stone et al., 2002).

For the RME test (Baron-Cohen et al., 2001), we translated the official version of the Baron-Cohen test (https://www.autismresearchcentre.com/tests/eyes-test-adult/) into Italian, since the first participants were tested before the Italian adaptation was published. Furthermore, we selected 30 out of the original 36 items, excluding items where the verbal label in Italian corresponded to a word with very low usage frequency. In order to ensure methodological consistency, data collection was carried out in the same manner for all the subsequent participants. An independent group of 15 healthy subjects validated the chosen stimuli. The selected 30 images were presented using PowerPoint on a 15″ screen. Each slide featured a black-and-white photograph of the eye region of a human face against a white background, accompanied by four adjectives (e.g., bothered, joking, passionate, comforting). Subjects are asked to choose the adjective that best describes the mental state expressed by the person in the image. There was no time limit for responding.

RMV (Rutherford et al., 2002; Golan et al., 2007) assesses the ability to recognize one's intention through the prosody. We adapted the original version of the task to Italian, including two different distractors for each trial (see below), selecting 35 new short sentences that were recorded by means of the Audacity software 1.2.6 (http://audacity.sourceforge.net/). Then, these sentences were validated by an independent group of healthy subjects and 30 of them were selected for the task.

Stimuli are presented both orally and in written form at the same time, on a computer screen through Microsoft Office PowerPoint. The items were designed so that the meaning of each sentence never matched the prosody with which it was pronounced. For instance, the sentence "I swear I have" typically indicates the completion of an action. However, when pronounced with a sarcastic tone, it implies the opposite, namely that the action has not been completed. The task required subjects to listen to the sentence and select the label that best describes the prosodic meaning conveyed by the sentence. The labels include: (i) an adjective that accurately reflects the prosody (correct answer); (ii) an adjective that corresponds to the sentence's semantic meaning (semantic error); (iii) an adjective that matches neither the prosody nor the meaning of the sentence (incorrect answer).

FP (Stone et al., 2002) assesses both the cognitive and affective components of ToM. Given the length of time required to administer the entire test, the 10 least complex stories in terms of comprehension were selected, choosing five stories that contain faux pas and five stories that do not (control tasks), from the Italian version developed by Massaro and colleagues (https://www.autismresearchcentre.com/tests/faux-pas-test-adult). The task requires participants to listen to a story read by an examiner. Some stories contain a faux pas, in which a character says or does something that unintentionally offends or embarrasses another character, while others do not. After each story, participants are asked to answer a series of questions designed to test their understanding of the social dynamics. These questions include: (i) Faux Pas detection; (ii) Theory of Mind questions; and (iii) Control questions to ensure basic comprehension. The total score was obtained from the sum of the single scores. Correct identification of faux pas and correct answers to ToM-related questions indicate an understanding of social nuances and the ability to infer others' mental states, while errors might indicate difficulties in recognizing or interpreting social information.

2.2.3 Neuropsychological assessment

All patients were submitted to a comprehensive neuropsychological evaluation. For the purpose of the present study, performances in the following tasks were considered: Mini-Mental State Examination (MMSE; Magni et al., 1996) for the assessment of the stage of the neurodegenerative disease, Benton test of facial recognition (Benton et al., 1983) to identify the presence of perceptual difficulties in processing faces and the Similarities subtest from the Italian revised version of the Wechsler Adult Intelligence Scale (Orsini and Laicardi, 1997), as a measure of executive functions.

2.3 Statistical analyses

Data management and analysis were performed using RStudio (2024; RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/) and Statistica (https://docs.tibco.com/).

Normality of variables was assessed using the Shapiro-Wilk test.

Parametric tests (repeated-measures ANOVAs; *t*-test for independent samples) and non-parametric tests (Kruskal-Wallis rank sum test; Pearson's Chi-Squared test) were used to investigate differences among groups. For *post-hoc* comparisons, Newman-Keuls tests were used and they were adjusted by Bonferroni corrections for multiple comparisons to account for the probability of committing type-1 errors. Finally, Pearson or Spearman's correlations were performed to assess the relationship between neuropsychological scores and emotion recognition test scores.

3 Results

Analyses of data distribution using the Shapiro-Wilk test showed that age and corrected MMSE scores were normally distributed, whereas education and gender distribution were not normal. Analyses of data distribution using the Shapiro-Wilk test showed that basic emotion recognition batteries, RME, RMV, Benton test of facial recognition and Similarities WAIS subtest were normally distributed, whereas FP and MMSE were not.

3.1 Demographic data

The one-way ANOVA was conducted with Age as within-subjects factor and Group (AD, FTD, HC) as between-subjects factor. The main effects of the Groups were not significant [$F_{(3,119)} = 2.11$, $p = 0.1$, $\eta^2p = 0.05$]. The Kruskal-Wallis ANOVA by Ranks sum test was conducted with Education as within-subject factor and Group (AD, FTD, HC) as between-subjects factor. The main effect of the Group was not significant ($H_3 = 3.81$, $p = 0.3$). The *T*-test for independent samples conducted on corrected MMSE scores between AD and FTD was not significant [$t_{(31)} = 0.06$, $p = 0.95$]. The Chi-squared test of independence was performed to analyze the distribution of gender across the groups (AD, FTD, HC). The test revealed a significant difference in gender distribution among the groups ($\chi^2 = 8.46$, df $= 3$, p-value $= 0.04$).

3.2 Basic emotion recognition batteries

3.2.1 Facial emotion recognition battery

A repeated measures ANOVA was conducted on the Facial Emotion Recognition Battery scores with Subtests (FM, FIR, FAN, FAS, and FAM) as within-subjects factor and Group (AD, FTD, HC) as between-subjects factor (Table 2). The main effects of Group [$F_{(2,100)} = 9.4$, $p < 0.001$, $\eta^2p = 0.16$] and Subtests [$F_{(4,400)} = 9.4$, $p < 0.001$, $\eta^2p = 0.78$] were statistically significant, as well as the interaction Subtests * Group [$F_{(8,400)} = 9.6$, $p < 0.001$, $\eta^2p = 0.16$]. The *post-hoc* comparisons showed (Table 2) that FTD patients significantly differed in the FAN ($M = 18.9$, SEM $= 0.6$; Figure 1, left) and in the FAM ($M = 15.8$, SEM $= 1.4$) subtests from both AD (FAN $M = 21.6$, SEM $= 0.1$, $p < 0.01$; FAM $M = 18.7$, SEM $= 0.2$, $p < 0.001$) and HC (FAN $M = 21.8$, SEM $= 0.3$, $p < 0.01$; FAM $M = 21.4$, SEM $= 0.4$, $p < 0.001$); in addition, AD differed from HC in the FAM subtest ($p < 0.01$). Interestingly, the significant differences resisted also when conducting an Analysis of Covariance (ANCOVA) considering age and education as covariates [$F_{(2,98)} = 14.2$, $p < 0.001$]. All the p's resisted Bonferroni's correction.

3.2.2 Emotional prosody recognition battery

A repeated measures ANOVA was conducted on the Emotional Prosody Recognition Battery scores with Subtests (VID, PrD, PrAN, PrAD) as within-subjects factor and Group (AD, FTD, HC) as between-subjects factor (Table 3). The main effects of Group [$F_{(2,100)} = 13$, $p < 0.001$, $\eta^2p = 0.21$] and of Subtests [$F_{(3,300)} = 1063.3$, $p < 0.001$, $\eta^2p = 0.91$] were significant. The interaction Subtests * Group was also significant [$F_{(6,300)} = 4.3$, $p < 0.001$, $\eta^2p = 0.08$]. The *post-hoc* analyses (Table 3) showed that FTD significantly differed in the PrAN ($M = 12.3$, SEM $= 1.2$; Figure 1, right) subtest from both AD ($M = 16.9$, SEM $= 0.2$, $p < 0.01$) and HC ($M = 18.2$, SEM $= 0.5$, $p < 0.001$). FTD patients' scores ($M = 33.6$, SEM $= 0.9$) were also significantly different from HC ($M = 37.7$, SEM $= 0.6$, $p < 0.001$) and AD ($M = 35.8$, SEM $= 0.2$, $p = 0.02$) in the PrAD subtest. The difference between FTD and AD did not resist Bonferroni's correction for multiple comparisons. All other p's resisted Bonferroni's correction. Interestingly, the significant differences resisted also when conducting an ANCOVA analysis considering age and education as covariates [$F_{(2,98)} = 14.7$, $p < 0.001$].

3.3 Social emotion recognition (ToM tasks)

One-way between-subject ANOVA on RME ratings revealed a significant effect of Group [$F_{(2,50)} = 4.2$, $p < 0.05$; $\eta^2p = 0.14$, Table 4]. *Post-hoc* comparisons (Table 4) that FTD ratings ($M = 14.9$, SEM $= 1$) were significantly lower compared to HC ($M = 19.1$, SEM $= 1.1$; $p < 0.05$) and to AD ($M = 17.9$, SEM $= 1$; $p < 0.05$). There was no significant difference between HC and AD (Figure 2, left).

One-way between-subject ANOVA on Semantic Errors of RMV revealed a significant effect of Group [$F_{(2,50)} = 10.4$, $p < 0.001$; $\eta^2p = 0.29$, Table 4]. *Post-hoc* comparisons (Table 4) revealed that FTD

TABLE 2 Descriptive statistics, repeated measures ANOVA and *post-hoc* results conducted on facial emotion recognition battery (FERB) scores.

	Descriptives			ANOVA			
	AD (n = 18)	FTD (n = 15)	HC (n = 70)	Group F	FERB F	Group × FERB F	Neuman Keuls *post-hoc* (Cohen's *d*)
				9.37*	**353.74***	**9.56***	
FM							
M	13.61	13.67	13.77				
(sd)	(0.85)	(0.62)	(0.68)				
FiR							
M	11.83	11.93	12.66				
(sd)	(1.65)	(2.46)	(2.09)				
FAN							
M	21.61	18.93	21.84				FTD < AD (1.18)***
(sd)	(2.28)	(2.25)	(2.64)				FTD < HC (1.19)***
FAS							
M	21.83	20.47	22.57				
(sd)	(2.31)	(3.25)	(2.53)				
FAM							
M	18.72	15.8	21.36				FTD < AD (0.6)***; AD < HC
(sd)	(4)	(5.53)	(3.11)				(1.14)***; FTD < HC (1.24)***

AD, Alzheimer Disease; FTD, Frontotemporal Dementia; HC, Healthy Controls; FM, Face Matching; FAN, Facial Affect Naming; FAS, Facial Affect Selection; FAM, Facial Affect Matching; M, mean; sd, standard deviation; ***p < 0.001.

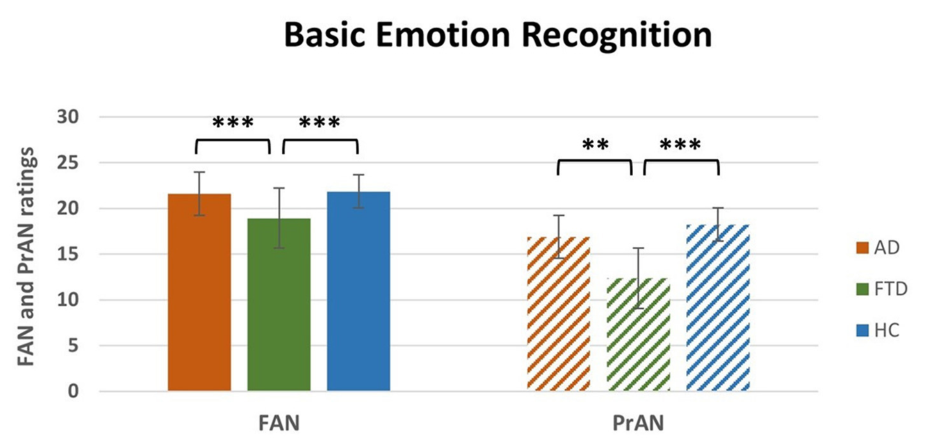


FIGURE 1
Interaction Subtest * Group *post-hoc* results for FAN and PrAN subtests. FAN, Facial Affect Naming; PrAN, Prosodic Affect Naming; AD, Alzheimer's Disease; FTD, FrontoTemporal Dementia; HC, Healthy Controls. **$p < 0.01$, ***$p < 0.001$.

ratings ($M = 11.2$, SEM = 0.6) were significantly higher compared to HC ($M = 6.6$, SEM = 0.8; $p < 0.001$) and AD ($M = 9$, SEM = 0.6; $p < 0.05$). Moreover, there was a significant difference between AD and HC ($p < 0.05$).

One-way between-subject ANOVA on RMV ratings revealed a significant effect of Group [$F_{(2,50)} = 17.3$, $p < 0.0001$; $\eta^2p = 0.41$, Table 4]. *Post-hoc* comparisons (Table 4) revealed that FTD ratings ($M = 11.6$, SEM = 1) were significantly lower compared to HC ($M = 19.6$, SEM = 1.1; $p < 0.001$) and to AD ($M = 16.6$, SEM = 0.7; $p < 0.001$). Moreover, there was a significant difference between AD and HC ($p < 0.05$; Figure 2, right).

Non-parametric one-way between-subject ANOVA (Kruskal-Wallis rank sum test) on FP ratings revealed a significant effect of Group ($H_2 = 19.1$, $p < 0.0001$, Table 4). *Post-hoc* comparisons (Table 4) revealed that FTD ($M = 16.9$, SEM = 3.1; $p < 0.001$) and AD ($M = 22.6$, SEM = 2.7; $p < 0.05$) ratings were significantly lower compared to HC ($M = 38.5$, SEM = 1.3). The difference between FTD and HC resisted Bonferroni's correction, whereas there was no significant difference between AD and FTD (Figure 3).

10.3389/fpsyg.2025.1535722

TABLE 3 Descriptive statistics, repeated measures ANOVA and *post-hoc* results conducted on emotional prosody recognition battery (EPRB) scores.

	Descriptives			ANOVA			
	AD ($n = 18$)	FTD ($n = 15$)	HC ($n = 70$)	Group F	EPRB F	Group × EPRB F	Neuman Keuls *post-hoc* (Cohen's *d*)
				12.99* **	**1063.32*** **	**4.26*** **	
VID							
M	12.17	10.87	12.01				
(sd)	(2.23)	(1.85)	(1.91)				
PrD							
M	13.06	11.87	13.41				
(sd)	(1.47)	(1.92)	(1.95)				
PrAN							
M	16.89	12.33	18.21				FTD < AD (1.14)***
(sd)	(3.20)	(4.65)	(4.18)				FTD < HC (1.33)***
PrAD							
M	35.83	33.60	37.69				FTD < HC (1)***
(sd)	(3.07)	(3.33)	(4.75)				FTD < AD (0.7)*

AD, Alzheimer Disease; FTD, Frontotemporal Dementia; HC, Healthy Controls; VID, Vocal Identity Discrimination; PrD, Prosodic Discrimination; PrAN, Prosodic Affect Naming; PrAD, Prosodic Affect Discrimination; M, mean; sd, standard deviation; *p < 0.05; ***p < 0.001.

TABLE 4 Descriptive statistics, repeated measures ANOVA and *post-hoc* results conducted on ToM scores.

	Descriptives			ANOVA	
	AD ($n = 18$)	FTD ($n = 15$)	HC ($n = 20$)	Group F	Neuman Keuls *post-hoc* (Cohen's *d*)
RMV					
M	16.61	11.6	19.6		FTD < AD (1.47)***; HC < FTD
(sd)	(2.97)	(3.79)	(4.83)	17.329***	(1.84)***; HC < AD (0.75)**
RME					
M	17.89	14.87	19.1		FTD < AD (0.77)*
(sd)	(4.06)	(3.81)	(4.95)	4.159*	HC < FTD (0.94)*
RMV-SE					
M	9	11.2	6.6		FTD < AD (0.91)*; HC < FTD
(sd)	(2.66)	(2.14)	(3.68)	10.400***	(1.53)***; HC < AD (0.75)*
				Group H (KW)	
FP					
M	44.39	39.27	54.9		HC < FTD (1.65)***
(sd)	(11.45)	(12.19)	(5.67)	19.1***	HC < AD (1.16)***

RMV, Reading the Mind from the Voice; RME, Reading the Mind from the Eyes; RMV-SE, Reading the Mind from the Voice-Semantic Errors; FP, Faux Pas; KW, Kruskal-Wallis;*p < 0.05; **p < 0.01; ***p < 0.001.

3.4 Neuropsychological tests

There was no significant difference between AD and FTD at the Benton test of facial recognition [$t_{(30.54)} = 1.38$, $p = 0.18$]. The mean score of Similarities WAIS subtest of FTD patients ($M = 27.32$, SEM = 3.8) was significantly different compared to the mean score of AD patients [$M = 39.91$, SEM = 1.5; $t_{(18.51)} = 3.06$, $p < 0.01$].

3.5 Correlations

Pearson's correlation analyses revealed three significant positive correlations: between Similarities WAIS subtest Test and RME in FTD patients ($r = 0.6$, $p < 0.05$); between Similarities WAIS subtest and RME in all patients ($r = 0.5$, $p < 0.01$); and between Similarities WAIS subtest and RMV in patients ($r = 0.4$, $p < 0.05$; Figure 4).

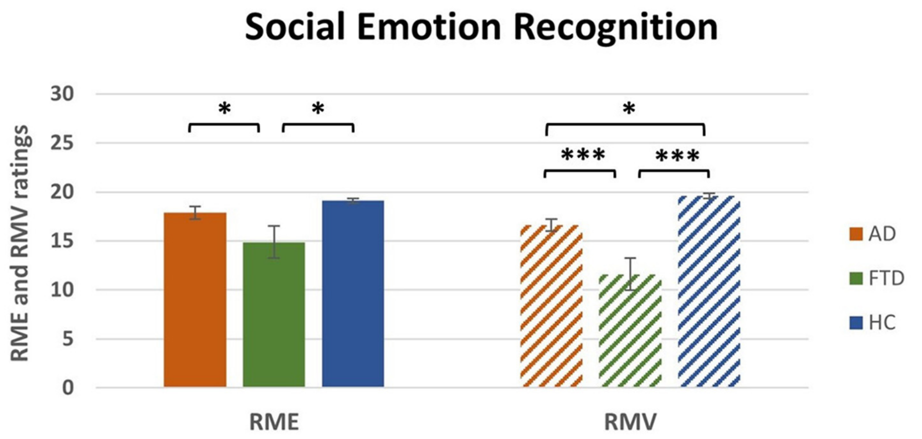


FIGURE 2
Main effect of group *post-hoc* results for RME and RMV tasks. RME, Reading the Mind from the Eyes; RMV, Reading the Mind in the Voice; AD, Alzheimer's Disease; FTD, FrontoTemporal Dementia; HC, Healthy Controls. $^{*}p < 0.05$, $^{***}p < 0.001$.

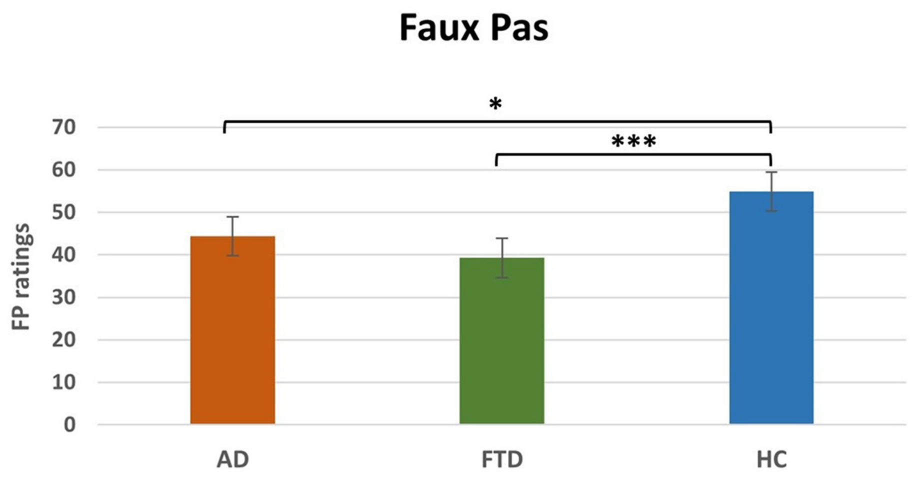


FIGURE 3
Main effect of group *post-hoc* results for FP. FP, Faux Pas; AD. Alzheimer's disease; FTD, FrontoTemporal Dementia; HC, Healthy Controls. $^{*}p < 0.05$, $^{***}p < 0.001$.

4 Discussion

In the present study we aimed to assess the social skills in the early stages of Alzheimer's Disease (AD) and FrontoTemporal Dementia (FTD), this information can be useful both for a better characterization and for a better clinical management of the two conditions. To this end, we recruited two groups of patients, one with AD and one with FTD, whose cognitive impairment was comparable (MMSE not significantly different between the two groups), and a control group of healthy participants. Both visual and auditory tasks were administered, both for basic emotion and for social emotion (ToM) recognition. In addition, to test cognitive ToM, the Faux Pas (FP) test was used. Overall, our results suggest that in early stages of FTD and AD there is an impairment of social cognition.

Regarding FTD, in line with the literature, our data show that these patients are significantly impaired, as compared to AD patients and HC, in all tasks that evaluate basic and social emotions processing. Specifically, FTD patients' deficits emerge, in comparison to AD patients, in those subtests of Facial Emotion Recognition Battery and Emotional Prosody Recognition Battery which require the association of an emotional (visual and auditory) expression with a verbal label, that is, in the Facial Affect Naming (FAN) and Prosodic Affect Naming (PrAN) subtests. Various studies showed consistent deficits in emotion recognition common to both visual (Bertoux et al., 2015; Jiskoot et al., 2021) and vocal stimuli (Dara et al., 2013; Wright et al., 2018), and particularly severe for negative emotions (Rosen et al., 2002; Fernandez-Duque and Black, 2005). Indeed, social cognition deficits are widely recognized as a hallmark of FTD, especially in the behavioral variant (bvFTD; Rascovsky et al., 2011). Notably, this impairment is particularly relevant in bvFTD as well as in the semantic variant of FTD (svFTD; Kumfor et al., 2014; Lee et al., 2020), and has been linked to structural atrophy in brain regions such as the

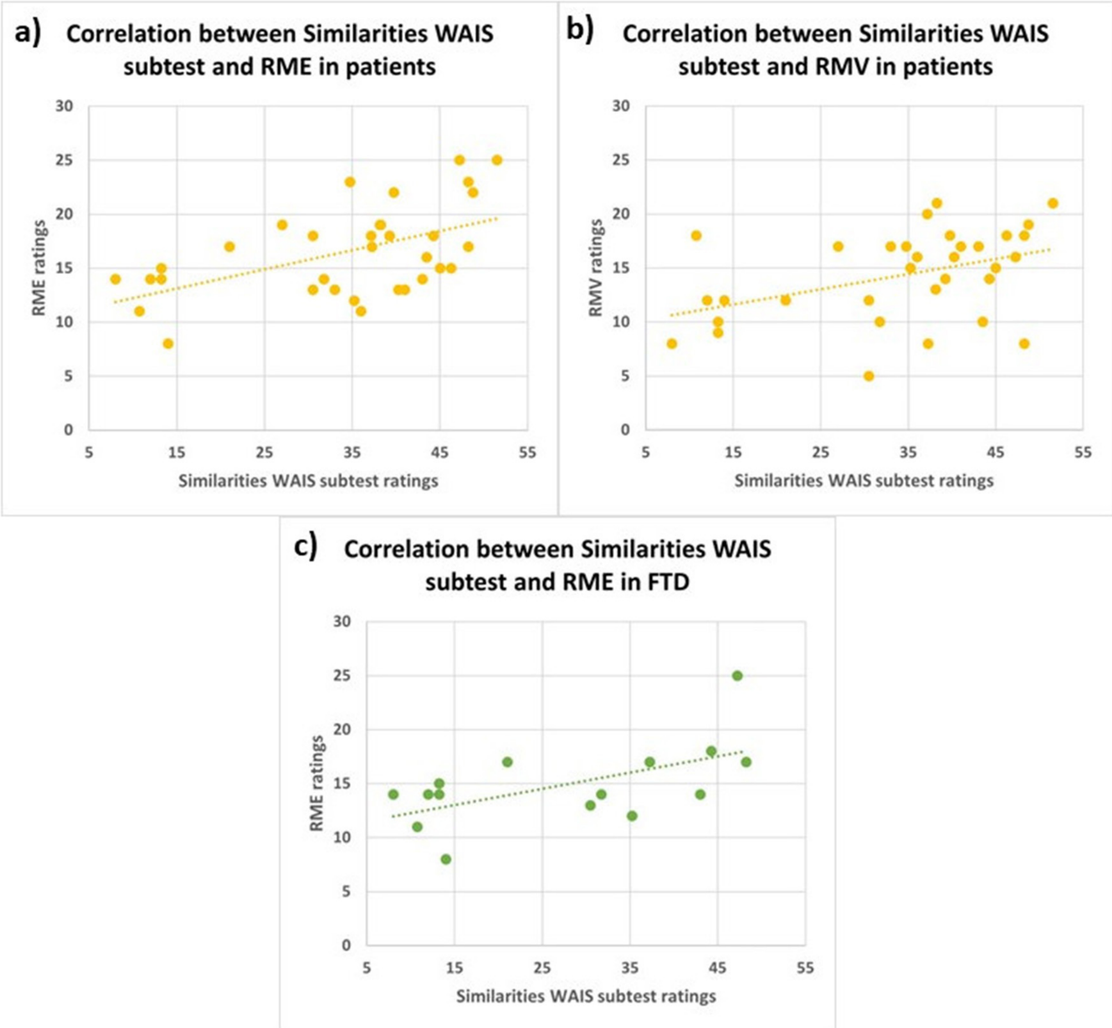


FIGURE 4
Correlations between similarities WAIS subtest and RME in all patients (**A**; $r = 0.5$, $p < 0.01$), between Similarities WAIS subtest and RMV in all patients (**B**; $r = 0.4$, $p < 0.05$) and between Similarities WAIS subtest and RME in FTD patients (**C**; $r = 0.6$, $p < 0.05$). RME, Reading the Mind in the Eyes; RMV, Reading the Mind in the Voice; FTD, FrontoTemporal Dementia.

anterior temporal lobes and the amygdala (Rosen et al., 2002, 2006; Kumfor et al., 2014; Lee et al., 2020). On the other hand, previous findings showed that tasks that increase the intensity of emotional expressions may mitigate recognition issues in bvFTD and primary progressive non-fluent aphasia (PNFA), suggesting that attentional and perceptual difficulties contribute to deficits in some FTD subtypes (Rascovsky et al., 2011). However, in svFTD, these issues are likely due to primary emotion processing impairments, rather than to cognitive overload. Interestingly, negative emotion recognition turned out to be particularly useful for differentiating FTD from AD (Bora et al., 2016), as AD patients typically show milder deficits.

Deficits in the ToM skills are usually found in FTD patients, especially in the behavioral variant, with respect to AD patients, both in visual (Gregory et al., 2002) and auditory modality (Orjuela-Rojas et al., 2021). Our findings demonstrate that FTD patients exhibit significantly greater impairments in both RME and RMV tests, compared to HC. Additionally, FTD patients show significantly worse performance in the RMV test compared to AD patients. Interestingly, especially in FTD patients, we found that Similarities WAIS subtest and RME scores positively correlated, showing that the higher the deficit in executive functions, the higher the impairment in recognizing emotions from eyes' cues. The association between executive functions and emotional recognition has been reported in healthy aging (Circelli et al., 2013) and in several psychiatric diseases (David et al., 2014; Yang et al., 2015; Williams et al., 2015), as well as in neurodegenerative diseases, such as Parkinson's Disease (Péron et al., 2012) and AD (Buçgün et al., 2023). In particular, social cognition skills could rely on executive processes, such as mental speed, cognitive flexibility, and inhibitory control to disregard personal viewpoints and concentrate on pertinent aspects, enabling the timely processing of all relevant information (David et al., 2014). Indeed, the ability to understand others' beliefs, intentions, and goals (cognitive ToM) relies on a frontotemporal network comprising the dorsolateral prefrontal cortex (dlPFC; cognitive processing of mental states and perspective-taking), the ventromedial prefrontal cortex (vmPFC), the orbitofrontal cortex (OFC), and the amygdala (processing and regulating emotional states; Amodio and Frith, 2006; Shamay-Tsoory and Aharon-Peretz, 2007; Abu-Akel and Shamay-Tsoory, 2011). On the other hand, the posterior superior temporal sulcus (pSTS), the temporo-parietal junction (TPJ), and medial prefrontal cortex (mPFC) belong to both emotional and cognitive ToM networks (Schurz and Perner, 2015; Molenberghs et al., 2016). This network facilitates representing others' mental states and differentiating them from one's own, regardless of the nature of the states.

Lastly, we found that FTD patients in the RMV test made a significant amount of semantic errors as compared to healthy controls. This indicates that these patients were able to understand the sentences' meaning, nevertheless they exhibited selective impairment in recognizing the affective aspects of prosody.

Focusing on AD, our study revealed a clear impairment in the recognition of emotional expression in the Facial Affect Matching (FAM) subtest, which requires to keep in memory and compare two emotional stimuli. This deficit could be explained by the overlap between areas engaged in memory tasks, and those involved in emotional processing, both prone to neurodegeneration in AD (Bediou et al., 2012). Indeed, most of the previous studies concluded that the ability to understand facial and prosodic emotional expressions is likely impaired because of the general cognitive decline observed in these patients (Amlerova et al., 2022; Buçgün et al., 2023). The mild deficits described in emotion recognition in the early stages of AD were more specifically related to low-intensity or negative emotions, such as sadness (Maki et al., 2013; Torres et al., 2015; Garcia-Cordero et al., 2021). On the other hand, emotion recognition seemed preserved in tasks with low cognitive demand (Luzzi et al., 2007). Furthermore, deficits in emotional processing in AD also extended across different sensory modalities (e.g., prosody; Amlerova et al., 2022). Therefore, the deficits observed in AD patients in the FAM subtest could be related to the high cognitive demand intrinsic to the task, since the patient is required to remember a face and associate it with an emotional label.

Regarding the ToM skills in AD patients, according to Wright et al. (2018), the recognition of affective prosody relies on a ventral processing stream involving the superior temporal cortex as well as the inferior and anterior temporal cortex in the right hemisphere. Impairments in this pathway may result in a compromised access to the Abstract Representations of Acoustic Characteristics that Convey Emotion (ARACCE; Wright et al., 2018). This is in line with our findings that AD patients are selectively impaired in recognizing emotions from the voice, that is, from the prosody, as compared to the HC group. These emotional recognition deficits in AD are consistent with neurodegeneration in temporal lobes (Bediou et al., 2012; Amlerova et al., 2022), affecting the abstract representations of acoustic features that convey emotions.

Finally, FTD and AD patients are significantly impaired in several Faux Pas (FP) subtests, including the affective and cognitive scores, as compared to the HC, whereas the groups of patients did not differ from each other. Our FP task contained several questions which enabled us to assess whether patients understood both the semantic aspects of the stories (control stories and questions) and the social gaffes (faux pas stories and questions).

Recent neuroimaging studies showed that the areas associated with the RME task are the left and right middle temporal gyri, superior temporal gyrus, cingulate gyrus, superior frontal gyrus, inferior frontal gyrus, middle frontal gyrus and left precentral gyrus. A recent FDG-PET and MRI (Magnetic Resonance Imaging) study hypothesizes that the ToM neural correlates can be categorized into hubs and spokes (Orso et al., 2022). Within the connectionist paradigm (van den Heuvel et al., 2009), it has been suggested that regions with greater connectivity to other components of a network (i.e., the "hubs") play a more crucial role in network functioning than those with less connectivity (i.e., the "spokes;" Hwang et al., 2013). Moreover, it has been hypothesized that damage to secondary nodes (spokes) can be compensated by the integrity of central nodes (hubs), whereas damage to the hubs themselves may result in clinical symptoms (van den Heuvel et al., 2009; Hwang et al., 2013). According to the structural connectivity and distribution of hypometabolism, hubs of the RME network were identified in frontal regions. This may explain ToM deficits commonly observed in FTD patients, where neurodegeneration impacts these hubs in the early stages of the disease (Adenzato et al., 2010). In contrast, in AD, their functional involvement typically becomes evident in the later stages of the disease, thus explaining the absence of ToM impairments in the early stages of the disease (Lucena et al., 2020). Indeed, our results in the RME subtest are consistent with these hypotheses, in that we only found deficits in FTD patients, which could likely be due to the neurodegeneration of these hubs.

To the best of our knowledge, this is the first extensive evaluation of emotional and social abilities in groups of neurodegenerative patients in the Italian population. Namely, we revised the RMV task to better assess the prosodic affective component of ToM in FTD patients. In particular, we introduced the possibility to evaluate semantic vs. non-semantic errors. The development of two tools for studying ToM abilities in the Italian language fills a gap in neuropsychological testing by providing instruments specifically adapted for use with Italian patients, which were previously unavailable. Furthermore, alongside the ability to quantify errors in the RMV test, we have introduced the capability to qualify these errors. This allows for the identification of patients with deficits in mental state processing stemming from semantic impairments, as opposed to those whose errors may be attributed to task complexity, thus reflecting the underlying neurodegenerative process. Thus, the modified versions of the two ToM tests are sensitive to detecting deficits that cannot be attributed to a generic neurodegenerative process. Combined with the two emotion processing batteries, they represent effective tools for both the quantification and qualification of social cognition impairments, even in patients with different neurological conditions such as traumatic brain injury, epilepsy (as demonstrated in our previous studies; Benuzzi et al., 2004; Bonora et al., 2011), focal lesions, and brain tumors.

Some limitations of the present study must be considered. Firstly, the involvement of a larger sample size for both groups of patients with dementia will be necessary. Specifically, in the group of patients with FTD, future studies should examine the impact of the different dementia variants on social abilities. Additionally, utilizing more and/or more refined tests than in the present study, with further tasks and tests, will enable a better understanding of the changes in the various components of ToM and emotion recognition in various forms of dementia.

Summing up, in the current study we found that FTD patients are significantly impaired in social cognition abilities, both in visual and in auditory modality, as compared to both AD patients and HC. On the other hand, in AD the emotional recognition impairment is prevalent in the auditory modality. Therefore, the introduction of the evaluation of these aspects in the clinical neuropsychological assessment could provide new insights into the cerebral localization of emotional and social skills, and into the neurodegenerative processes that may affect them.

Considering the clinical and social influence of social cognition impairments in these two neurodegenerative diseases, this study aimed to provide a more comprehensive characterization of their impact. Specifically, we employed an extensive assessment protocol designed to evaluate both visual and auditory processing across basic and social emotions within the same patient groups. While these tests may not produce clear differences sufficient for individual diagnosis, our objective is to offer a robust and reliable framework for delineating the behavioral and emotional profiles characteristic of AD and FTD. This, in turn, can serve as a valuable tool for enhancing our understanding of these diseases and facilitating improved clinical management strategies, including tailored therapeutic interventions and caregiver support.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Comitato Etico di Ateneo per la Ricerca, CEAR; Prot. n 83243. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CS: Investigation, Methodology, Validation, Writing – review & editing. VZ: Formal analysis, Visualization, Writing – original draft, Writing – review & editing. CC: Formal analysis, Validation, Writing – original draft, Writing – review & editing. MM: Conceptualization, Methodology, Supervision, Writing – review & editing, Investigation. FR: Writing – original draft, Writing – review & editing. OC: Software, Writing – review & editing. MT: Writing – review & editing, Investigation. FL: Conceptualization, Project administration, Supervision, Writing – review & editing. PN: Conceptualization, Supervision, Writing – review & editing. FB: Conceptualization, Data curation, Investigation, Project administration, Supervision, Validation, Writing – review & editing, Methodology, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Abbruzzese, M., Bellodi, L., Ferri, S., and Scarone, S. (1995). Frontal lobe dysfunction in schizophrenia and obsessive-compulsive disorder: a neuropsychological study. *Brain Cogn.* 27, 202–212. doi: 10.1006/brcg.1995.1017

Abu-Akel, A., and Shamay-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia* 49, 2971–2984. doi: 10.1016/j.neuropsychologia.2011.07.012

Adenzato, M., Cavallo, M., and Enrici, I. (2010). Theory of mind ability in the behavioural variant of frontotemporal dementia: an analysis of the neural, cognitive, and social levels. *Neuropsychologia* 48, 2–12. doi: 10.1016/j.neuropsychologia.2009.08.001

Adolphs, R. (1999). Social cognition and the human brain. *Trends Cogn. Sci.* 3, 469–479. doi: 10.1016/S1364-6613(99)01399-6

Adolphs, R., Tranel, D., and Damasio, A. R. (1998). The human amygdala in social judgment. *Nature* 393, 470–474. doi: 10.1038/30982

Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372, 669–672. doi: 10.1038/372669a0

Amlerova, J., Laczó, J., Nedelska, Z., Laczó, M., Vyhnálek, M., Zhang, B., et al. (2022). Emotional prosody recognition is impaired in Alzheimer's disease. *Alzheimer. Res. Ther.* 14:50. doi: 10.1186/s13195-022-00989-7

Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277. doi: 10.1038/nrn1884

Ariatti, A., Benuzzi, F., and Nichelli, P. (2008). Recognition of emotions from visual and prosodic cues in Parkinson's disease. *Neurol. Sci.* 29, 219–227. doi: 10.1007/s10072-008-0971-9

Baillargeon, R., Scott, R. M., and He, Z. (2010). False-belief understanding in infants. *Trends Cogn. Sci.* 14, 110–118. doi: 10.1016/j.tics.2009.12.006

Baron-Cohen, S. (1997). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge: MIT Press.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., and Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *J. Autism Dev. Disord.* 29, 407–418. doi: 10.1023/A:1023035012436

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry Allied Discipl.* 42, 241–251. doi: 10.1111/1469-7610.00715

Bediou, B., Brunelin, J., d'Amato, T., Fecteau, S., Saoud, M., Hénaff, M. A., et al. (2012). A comparison of facial emotion processing in neurological and psychiatric conditions. *Front. Psychol.* 3:98. doi: 10.3389/fpsyg.2012.00098

Benton, A., Hamsher, K. D., Varney, N., and Spreen, O. (1983). *Contribution to Neuropsychological Assessment*. New York, NY: Oxford University Press.

Benuzzi, F., Meletti, S., Zamboni, G., Calandra-Buonaura, G., Serafini, M., Lui, F., et al. (2004). Impaired fear processing in right mesial temporal sclerosis: a fMRI study. *Brain Res. Bull.* 63, 269–281. doi: 10.1016/j.brainresbull.2004.03.005

Bertoux, M., de Souza, L. C., Sarazin, M., Funkiewiez, A., Dubois, B., and Hornberger, M. (2015). How preserved is emotion recognition in Alzheimer disease compared with behavioral variant frontotemporal dementia? *Alzheimer Dis. Assoc. Disord.* 29, 154–157. doi: 10.1097/WAD.0000000000000023

Bonora, A., Benuzzi, F., Monti, G., Mirandola, L., Pugnaghi, M., Nichelli, P., et al. (2011). Recognition of emotions from faces and voices in medial temporal lobe epilepsy. *Epilepsy Behav.* 20, 648–654. doi: 10.1016/j.yebeh.2011.01.027

Bora, E., Velakoulis, D., and Walterfang, M. (2016). Meta-analysis of facial emotion recognition in behavioral variant frontotemporal dementia: comparison with Alzheimer disease and healthy controls. *J. Geriatric Psychiatry Neurol.* 29, 205–211. doi: 10.1177/0891988716640375

Bouhuys, A. L., Geerts, E., and Gordijn, M. C. (1999). Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse: a longitudinal study. *J. Nervous Mental Dis.* 187, 595–602. doi: 10.1097/00005053-199910000-00002

Buçgün, I., Korkmaz, S. A., and Öyekçin, D. G. (2023). Facial emotion recognition is associated with executive functions and depression scores, but not staging of dementia, in mild-to-moderate Alzheimer's disease. *Brain Behav.* 14:e3390. doi: 10.1002/brb3.3390

Castelli, I., Pini, A., Alberoni, M., Liverta-Sempio, O., Baglio, F., Massaro, D., et al. (2011). Mapping levels of theory of mind in Alzheimer's disease: a preliminary study. *Aging Mental Health* 15, 157–168. doi: 10.1080/13607863.2010.513038

Ceravolo, L., Frühholz, S., Pierce, J., Grandjean, D., and Péron, J. (2021). Basal ganglia and cerebellum contributions to vocal emotion processing as revealed by high-resolution fMRI. *Sci. Rep.* 11:10645. doi: 10.1038/s41598-021-90222-6

Chaudhary, S., Zhornitsky, S., Chao, H. H., van Dyck, C. H., and Li, C. R. (2022). Emotion processing dysfunction in Alzheimer's disease: an overview of behavioral findings, systems neural correlates, and underlying neural biology. *Am. J. Alzheimer. Dis. Other Dement.* 37, 1–33. doi: 10.1177/15333175221082834

Circelli, K. S., Clark, U. S., and Cronin-Golomb, A. (2013). Visual scanning patterns and executive function in relation to facial emotion recognition in aging. *Neuropsychol. Dev. Cogn. Sect. B Aging Neuropsychol. Cogn.* 20, 148–173. doi: 10.1080/13825585.2012.675427

Craig, A. D. (2009). How do you feel–now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555

Dara, C., Kirsch-Darrow, L., Ochfeld, E., Slenz, J., Agranovich, A., Vasconcellos-Faria, A., et al. (2013). Impaired emotion processing from vocal and facial cues in frontotemporal dementia compared to right hemisphere stroke. *Neurocase* 19, 521–529. doi: 10.1080/13554794.2012.701641

David, D. P., Soeiro-de-Souza, M. G., Moreno, R. A., and Bio, D. S. (2014). Facial emotion recognition and its correlation with executive functions in bipolar I patients and healthy controls. *J. Affect. Disord.* 152–154, 288–294. doi: 10.1016/j.jad.2013.09.027

de Lucena, A. T., Dos Santos, T. T. B. A., Santos, P. F. A., and Dourado, M. C. N. (2023). Affective theory of mind in people with mild and moderate Alzheimer's disease. *Int. J. Geriatric Psychiatry* 38:e6032. doi: 10.1002/gps.6032

Demichelis, O. P., Coundouris, S. P., Grainger, S. A., and Henry, J. D. (2020). Empathy and theory of mind in Alzheimer's disease: a meta-analysis. *J. Int. Neuropsychol. Soc.* 26, 963–977. doi: 10.1017/S1355617720000478

Dodich, A., Crespi, C., Santi, G. C., Cappa, S. F., and Cerami, C. (2021). Evaluation of discriminative detection abilities of social cognition measures for the diagnosis of the behavioral variant of frontotemporal dementia: a systematic review. *Neuropsychol. Rev.* 31, 251–266. doi: 10.1007/s11065-020-09457-1

Ekman, P. (1999). "Basic emotions," in *Handbook of Cognition and Emotion*, eds. T. Dalgleish and M. J. Power (New York, NY: John Wiley and Sons Ltd), 45–60. doi: 10.1002/0470013494.ch3

Fernandez-Duque, D., and Black, S. E. (2005). Impaired recognition of negative facial emotions in patients with frontotemporal dementia. *Neuropsychologia* 43, 1673–1687. doi: 10.1016/j.neuropsychologia.2005.01.005

Frith, C. D. (2008). Social cognition. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 363, 2033–2039. doi: 10.1098/rstb.2008.0005

Garcia-Cordero, I., Migeot, J., Fittipaldi, S., Aquino, A., Campo, C. G., García, A., et al. (2021). Metacognition of emotion recognition across neurodegenerative diseases. *Cortex* 137, 93–107. doi: 10.1016/j.cortex.2020.12.023

Golan, O., Baron-Cohen, S., Hill, J. J., and Rutherford, M. D. (2007). The 'reading the mind in the voice' test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *J. Autism Dev. Disord.* 37, 1096–1106. doi: 10.1007/s10803-006-0252-5

Gossink, F., Schouws, S., Krudop, W., Scheltens, P., Stek, M., Pijnenburg, Y., et al. (2018). Social cognition differentiates behavioral variant frontotemporal dementia from other neurodegenerative diseases and psychiatric disorders. *Am. J. Geriatric Psychiatry* 26, 569–579. doi: 10.1016/j.jagp.2017.12.008

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.* 8, 145–146. doi: 10.1038/nn1392

Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., et al. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain* 125, 752–764. doi: 10.1093/brain/awf079

Gur, R. C., Erwin, R. J., Gur, R. E., Zwil, A. S., Heimberg, C., and Kraemer, H. C. (1992). Facial emotion discrimination: II. Behavioral findings in depression. *Psychiatry Res.* 42, 241–251. doi: 10.1016/0165-1781(92)90116-K

Hwang, K., Hallquist, M. N., and Luna, B. (2013). The development of hub architecture in the human functional brain network. *Cereb. Cortex* 23, 2380–2393. doi: 10.1093/cercor/bhs227

Jiskoot, L. C., Poos, J. M., Vollebergh, M. E., Franzen, S., van Hemmen, J., Papma, J. M., et al. (2021). Emotion recognition of morphed facial expressions in presymptomatic and symptomatic frontotemporal dementia, and Alzheimer's dementia. *J. Neurol.* 268, 102–113. doi: 10.1007/s00415-020-10096-y

Johnstone, T., van Reekum, C. M., Oakes, T. R., and Davidson, R. J. (2006). The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. *Soc. Cogn. Affect. Neurosci.* 1, 242–249. doi: 10.1093/scan/nsl027

Kanwisher, N., and Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 361, 2109–2128. doi: 10.1098/rstb.2006.1934

Kessels, R. P. C., Waanders-Oude Elferink, M., and van Tilborg, I. (2021). Social cognition and social functioning in patients with amnestic mild cognitive impairment or Alzheimer's dementia. *J. Neuropsychol.* 15, 186–203. doi: 10.1111/jnp.12223

Kumfor, F., Honan, C., McDonald, S., Hazelton, J. L., Hodges, J. R., and Piguet, O. (2017). Assessing the "social brain" in dementia: applying TASIT-S. *Cortex* 93, 166–177. doi: 10.1016/j.cortex.2017.05.022

Kumfor, F., Sapey-Triomphe, L. A., Leyton, C. E., Burrell, J. R., Hodges, J. R., and Piguet, O. (2014). Neuroanatomical correlates of emotion recognition deficits in behavioral variant frontotemporal dementia: a voxel-based morphometry study. *PLoS ONE* 8:e67457. doi: 10.1371/journal.pone.0067457

Lee, J. H., Jeong, H. K., Kim, Y. H., et al. (2020). Facial emotion recognition and its neuroanatomical correlates in elderly Korean adults with mild cognitive impairment, Alzheimer's disease, and frontotemporal dementia. *Front. Psychol.* 11:571760. doi: 10.3389/fnagi.2019.00091

Lucena, A. T., Bhalla, R. K., Belfort Almeida Dos Santos, T. T., and Dourado, M. C. N. (2020). The relationship between theory of mind and cognition in Alzheimer's disease: a systematic review. *J. Clin. Exp. Neuropsychol.* 42, 223–239. doi: 10.1080/13803395.2019.1710112

Luzzi, S., Piccirilli, M., and Provinciali, L. (2007). Perception of emotions on happy/sad chimeric faces in Alzheimer disease: relationship with cognitive functions. *Alzheimer Dis. Assoc. Disord.* 21, 130–135. doi: 10.1097/WAD.0b013e318064f445

Magni, E., Binetti, G., Bianchetti, A., Rozzini, R., and Trabucchi, M. (1996). Mini-mental state examination: a normative study in Italian elderly population. *Euro. J. Neurol.* 3, 198–202. doi: 10.1111/j.1468-1331.1996.tb00423.x

Maki, Y., Yoshida, H., Yamaguchi, T., and Yamaguchi, H. (2013). Relative preservation of the recognition of positive facial expression "happiness" in Alzheimer disease. *Int. Psychogeriatr.* 25, 105–110. doi: 10.1017/S1041610212001482

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer. Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005

Molenberghs, P., Johnson, H., Henry, J. D., and Mattingley, J. B. (2016). Understanding the minds of others: a neuroimaging meta-analysis. *Neurosci. Biobehav. Rev.* 65, 276–291. doi: 10.1016/j.neubiorev.2016.03.020

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

Orjuela-Rojas, J. M., Montañés, P., Rodríguez, I. L. L., and González-Marín, N. R. (2021). Recognition of musical emotions in the behavioral variant of frontotemporal dementia. *Rev. Colomb. Psiquiatr.* 50, 74–81. doi: 10.1016/j.rcp.2020.01.002

Orsini, A., and Laicardi, C. (1997). *WAIS-R. Contributo alla taratura italiana.* Firenze: OS Organizzazioni Speciali.

Orso, B., Lorenzini, L., Arnaldi, D., Girtler, N., Brugnolo, A., Doglione, E., et al. (2022). The role of hub and spoke regions in theory of mind in early Alzheimer's disease and frontotemporal dementia. *Biomedicines* 10:544. doi: 10.3390/biomedicines10030544

Perner, J., and Wimmer, H. (1985). John thinks what Mary thinks that...: attribution of second order false belief by 5- to 10-years old children. *J. Exp. Child Psychol.* 39, 437–471. doi: 10.1016/0022-0965(85)90051-7

Péron, J., Dondaine, T., Le Jeune, F., Grandjean, D., and Vérin, M. (2012). Emotional processing in Parkinson's disease: a systematic review. *Mov. Disord.* 27, 186–199. doi: 10.1002/mds.24025

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., et al. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 134(Pt 9), 2456–2477. doi: 10.1093/brain/awr179

Rosen, H. J., Perry, R. J., Murphy, J., Kramer, J. H., Mychack, P., Schuff, N., et al. (2002). Emotion comprehension in the temporal variant of frontotemporal dementia. *Brain* 125(Pt 10), 2286–2295. doi: 10.1093/brain/awf225

Rosen, H. J., Wilson, M. R., Schauer, G. F., Allison, S., Gorno-Tempini, M. L., Pace-Savitsky, C., et al. (2006). Neuroanatomical correlates of impaired recognition of emotion in dementia. *Neuropsychologia* 44, 365–373. doi: 10.1016/j.neuropsychologia.2005.06.012

Rutherford, M. D., Baron-Cohen, S., and Wheelwright, S. (2002). Reading the mind in the voice: a study with normal adults and adults with Asperger syndrome and high functioning autism. *J. Autism Dev. Disord.* 32, 189–194. doi: 10.1023/A:1015497629971

Sanavio, E., Bertolotti, G., Michielin, P., Vidotto, G., and Zotti, A. M. (1986). *CBA-2.0 Scale Primarie: Manuale. Una batteria a largo spettro per l'assessment psicologico.* Firenze: GIUNTI – Organizzazioni Speciali.

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage* 28, 848–858. doi: 10.1016/j.neuroimage.2005.06.023

Schirmer, A., and Adolphs, R. (2017). Emotion perception from face, voice, and touch: comparisons and convergence. *Trends Cogn. Sci.* 21, 216–228. doi: 10.1016/j.tics.2017.01.001

Schurz, M., and Perner, J. (2015). An evaluation of neurocognitive models of theory of mind. *Front. Psychol.* 6:1610. doi: 10.3389/fpsyg.2015.01610

Shamay-Tsoory, S. G., and Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia* 45, 3054–3067. doi: 10.1016/j.neuropsychologia.2007.05.021

Sica, C., and Ghisi, M. (2007). "The Italian versions of the beck anxiety inventory and the beck depression inventory-II: psychometric properties and discriminant power," in *Leading-Edge Psychological Tests and Testing Research*, ed. M. A. Lange (Hauppauge, NY: Nova Science Publishers), 27–50.

Stone, V., Baron-Cohen, S., Knight, R. T., Gregory, C., Lough, S., Erzinclioglu, S., et al. (2002). Faux Pas recognition test (adult version). *J. Cogn. Neurosci.* 10, 640–656. doi: 10.1162/089892998562942

Todorov, A., Mende-Siedlecki, P., and Dotsch, R. (2013). Social judgments from faces. *Curr. Opin. Neurobiol.* 23, 373–380. doi: 10.1016/j.conb.2012.12.010

Torres, B., Santos, R. L., Sousa, M. F., Simões Neto, J. P., Nogueira, M. M., Belfort, T. T., et al. (2015). Facial expression recognition in Alzheimer's disease: a longitudinal study. *Arq. Neuropsiquiatr.* 73, 383–389. doi: 10.1590/0004-282X20150009

van den Heuvel, M. P., Stam, C. J., Kahn, R. S., and Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *J. Neurosci.* 29, 7619–7624. doi: 10.1523/JNEUROSCI.1443-09.2009

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664. doi: 10.1016/S0896-6273(03)00679-2

Williams, G. E., Daros, A. R., Graves, B., McMain, S. F., Links, P. S., and Ruocco, A. C. (2015). Executive functions and social cognition in highly lethal self-injuring patients with borderline personality disorder. *Pers. Disord.* 6, 107–116. doi: 10.1037/per0000105

Wright, A., Saxena, S., Sheppard, S. M., and Hillis, A. E. (2018). Selective impairments in components of affective prosody in neurologically impaired individuals. *Brain Cogn.* 124, 29–36. doi: 10.1016/j.bandc.2018.04.001

Yang, C., Zhang, T., Li, Z., Heeramun-Aubeeluck, A., Liu, N., Huang, N., et al. (2015). The relationship between facial emotion recognition and executive functions in first-episode patients with schizophrenia and their siblings. *BMC Psychiatry* 15:241. doi: 10.1186/s12888-015-0618-3

Zhou, P., Ma, H., Zou, B., Zhang, X., Zhao, S., Lin, Y., et al. (2023). A conceptual framework of cognitive-affective theory of mind: towards a precision identification of mental disorders. *NPJ Mental Health Res.* 2:12. doi: 10.1038/s44184-023-00031-0

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 06 March 2025
DOI 10.3389/fpsyg.2025.1525368

OPEN ACCESS

EDITED BY
Ilaria Gabbatore,
University of Turin, Italy

REVIEWED BY
Roberta Fadda,
University of Cagliari, Italy
Schnell Zsuzsanna,
University of Pécs, Hungary

*CORRESPONDENCE
Alessia Cornaggia
✉ cornaggia13@gmail.com

†These authors share first authorship

RECEIVED 09 November 2024
ACCEPTED 14 February 2025
PUBLISHED 06 March 2025

CITATION
Bianco F, Cornaggia A, Massaro D, Marchetti A and Castelli I (2025) Desires and beliefs: the development of second-order Theory of Mind reasoning in preschoolers and in school-age children.
Front. Psychol. 16:1525368.
doi: 10.3389/fpsyg.2025.1525368

Desires and beliefs: the development of second-order Theory of Mind reasoning in preschoolers and in school-age children

Federica Bianco[1†], Alessia Cornaggia[1*†], Davide Massaro[2], Antonella Marchetti[2] and Ilaria Castelli[1†]

[1]Department of Human and Social Sciences, University of Bergamo, Bergamo, Italy, [2]Department of Psychology, Catholic University of Sacred Heart, Milan, Italy

Introduction: Theory of Mind development is crucial for social life. Most studies on the development of this skill have focused on first-order recursive thinking, while the transition to second-order thinking remains relatively unexplored.

Methods: To address this gap, we administered a novel second-order Theory of Mind task to 59 children between the ages of 5 and 8 years. This task manipulated desires (desire to obtain, "positive desire," vs. desire to avoid, "negative desire") and beliefs (true vs. false) based on previous studies of first-order scaling.

Results: Results indicate that the tasks involving positive desire seem to be easier than negative counterparts, and that the tasks involving true belief are easier than those involving false belief. All children performed below chance level in negative desire and in false belief conditions, while only older participants performed above chance level in true belief – positive desire condition. There was also a significant main effect favoring positive desire and true belief.

Discussion: Our findings provide preliminary evidence for the developmental acquisitions of second-order recursive thinking about the understanding of desires and beliefs.

KEYWORDS

Theory of Mind, second-order reasoning, desires, beliefs, scaling

1 Introduction

Early studies in Theory of Mind (ToM, Premack and Woodruff, 1978; Wimmer and Perner, 1983) have started a large and complex body of research concerning the development of the understanding of specific mental states such as intentions, desires, and beliefs (Apperly et al., 2011), and how this competence is interconnected with other developmental domains (Coull et al., 2006). Indeed, ToM is first of all the ability to recognize the presence of thoughts and feelings in one's own and other's minds, but it also represents the possibility to reason about these contents and about how they are associated with behavior and the responses to the context's influences (Lieberman, 2007).

While previous research in the ToM domain has provided a deep understanding of first-order reasoning (i.e., "*I think that you think*"), it has left areas of discontinuity in the study of higher-order ToM development (Apperly et al., 2011; Peterson and Wellman, 2019; Wellman, 2012). The transition from first- to second-order reasoning (i.e., *"I think that you think that he/she thinks"*), has emerged as the main point of such discontinuity. Previous literature has not provided a precise sequence of acquisition for understanding different kinds of mental states, such as emotions, desires, and beliefs in the second-order reasoning (Apperly et al., 2011; Osterhaus and Bosacki, 2022). In this work, we aim to understand the processes that lead the child to master the second-order false belief task to advance our knowledge of ToM performance in the *"uncharted waters of middle childhood"* (Hughes, 2016, p. 4).

1.1 First-order scaling

To assess first-order ToM acquisition in preschoolers, Wellman and Liu (2004) developed a five-item ToM scale which aims to describe, rather than to explain, the subjective understanding of different mental states, in the first-order domain. The results indicated a progression: the first task to be overcome seems to be the diverse desire, followed, respectively, by diverse belief, knowledge access, false belief, and hidden emotions (Wellman and Liu, 2004). Furthermore, new acquisitions were not simply added to previous ones, but rather the first achievements mediated the understanding of more complex mental states.

Subsequently, other studies (Peterson et al., 2012; Rivas-Garcia et al., 2020) have used the ToM scale proposed by Wellman and Liu (2004) to further investigate ToM development. These studies have introduced variations to the original measurement tool, specifically, they focused on beliefs and emotions, including diverse desires as a single task at the basis of all other achievements.

The belief-desire reasoning in the first-order domain was also investigated by Apperly et al. (2011) in a sample of children aged between 6 and 11 years. The authors showed that even younger children made fewer errors, and responded faster to true belief (when reality and beliefs coincide) and to positive desires (when a person wants something) compared to false belief (when reality and beliefs conflict) and to negative desires (when a person wants to avoid something) (Apperly et al., 2011). The pattern of errors and response times confirmed that the most challenging conditions were those involving reasoning about false beliefs and negative desires, not only for the children but also for the adults (Apperly et al., 2011). The results of this pivotal study showed a developmental progression from true belief to false belief and from positive desire to negative desire that was consistent across age groups. Additionally, older children outperformed younger children (Apperly et al., 2011).

1.2 The continuity in development from first-order to advanced ToM reasoning

The term advanced ToM refers to all the developmental acquisitions in understanding the mind and reasoning about mental states that occur after mastery of first-order reasoning (Miller, 2022). Its critical developmental period is between 6 and 10 years (Hughes and Devine, 2015) and continues throughout the life-span (Miller, 2009). Peterson and Wellman (2019) conducted a longitudinal study with children aged three to thirteen, exploring the development of ToM in middle childhood. The initial level of ToM understanding was found to be the best predictor of ToM performance in older children (Peterson and Wellman, 2019). The transition from early to advanced ToM abilities could be represented by the achievement of second-order reasoning, providing a link and continuity between the preschool years and middle childhood. Some research suggests that second-order false belief reasoning begins to emerge around the age of 5 or 6 (Miller, 2009). By the age of 7, success rates reach approximately 65%, and typically developing children complete second-order false belief tasks with 100% accuracy by the age of 11 (Arslan et al., 2013). Some studies of second-order reasoning have compared the traditional task proposed by Perner and Wimmer (1985) with the more simplified version of Sullivan et al. (1994). These two different measures placed the age of emergence at different points: 7 years for Perner and Wimmer (1985), and 5 years for Sullivan et al. (1994). The two tasks differed in the number of characters and scenes involved, the length of the stories, and the feedback provided for probe questions. Furthermore, Sullivan and colleagues included a second-order ignorance question that may help children understand false belief 2 years earlier (Hogrefe et al., 1986). These task characteristics appear to help mitigate the processing demands that might interfere with the detection of second-order false belief reasoning in children (Coull et al., 2006; Sullivan et al., 1994). Indeed, second-order tasks require not only a more sophisticated level of ToM, but also greater memory and language skills than first-order tasks (Miller, 2022).

Arslan et al. (2017) investigated the possible relationship between the ability to solve first and second-order ToM tasks, in the form of stories, in children aged 5 to 6 years. They used an instance-based learning model and found that failure on the second-order tasks was associated with answers based on first-order reasoning (Arslan et al., 2017). 17% of the sample answered the second-order tasks correctly, and the majority of incorrect responses appeared to be due to the influence of first-order reasoning, which seemed to interfere with second-order reasoning (Arslan et al., 2017). In a later training study (Arslan et al., 2020) with 5-year-old children, it appeared that the failure to perform second-order tasks was also due to a lack of experience with this type of reasoning and its justification. In recent years, various training studies (Bianco et al., 2016, 2019, 2021; Lecce et al., 2014; Lombardi et al., 2022; for a review see Bianco and Castelli, 2023) have demonstrated the possibility of improving second-order reasoning in middle childhood. They have also provided valuable insights into the continuity of ToM acquisitions from first-order to second-order and advanced ToM, identifying the same developmental engine of maturation in mental-state conversations. Furthermore, Bianco et al.' (2021) training study found that the age range of 7 to 8 years is a sensitive period for achieving second-order ToM reasoning.

As proposed by Osterhaus et al. (2016), the literature suggests various methods to assess different components of advanced ToM. However, there has been a lack of understanding about the progressive and continuous development of this ability, which is crucial for social experience. Indeed, there is no clear and systematic evidence on the development of ToM abilities after or at the highest steps of Wellman and Liu ToM Scale (Osterhaus et al., 2016). To be exhaustive, there is evidence of age effects on different types of tasks, but what is missing is a clear framework for the intra-individual development of the various components of advanced ToM knowledge (Miller, 2022). To the best of our knowledge, the first attempt in this direction was the study conducted by Osterhaus et al. (2016). Children compiled a scale of 24 Advanced ToM items, which were grouped into three factors: social reasoning, reasoning about ambiguity, and recognizing transgressions of social norms (Osterhaus et al., 2016). The three-factor structure was then found to be valid for both children aged 8–10 and younger children aged 5–8 (Osterhaus and Koerber, 2021a). A weak correlation was found between first-order and advanced ToM, suggesting conceptual continuity but also highlighting the difference between the two constructs (Osterhaus and Koerber, 2021a). Furthermore, Osterhaus and Koerber (2021b) in a longitudinal study showed that ToM development from 5 to 10 years was non-linear, with a plateau phase after the age of 7 years. This work also suggests that second-order reasoning could be considered one of the first mechanisms of advanced ToM, linking these mature forms of ToM to the previous competence, namely first-order false belief reasoning (Osterhaus and Koerber, 2021b).

1.3 ToM understanding executive functioning and verbal ability

The literature on classical first-order ToM has established associations with language (Astington and Baird, 2012; Belacchi, 2022; Milligan et al., 2007) and executive functions (Doenyas et al., 2018; Traverso et al., 2022).

The literature now considers language as a tool for conveying knowledge not only about the external world, but also in the representation of internal states (Belacchi, 2022) and there are many hypotheses about its relationship with ToM (Harris et al., 2005; Lockl and Schneider, 2007). The components of language permit the comprehension of the multiplicity of representations of reality and thus to consider ourselves and others as mental agents (Belacchi, 2022).

The relationship between ToM and executive function has been deeply studied and discussed in literature (Osterhaus and Bosacki, 2022; Traverso et al., 2022). Executive functions are a set of skills that allow individuals to anticipate, plan, set goals, implement projects, and monitor/modify the behavior to adapt to new conditions (Razza and Blair, 2009; Traverso et al., 2022). As Apperly et al. (2011) pointed out, executive demands could influence the interaction between beliefs and desires, specifically increasing the demand for inhibitory control when false beliefs and negative desire were combined in the same task. Even data from adults suggest a correlation between measures of processing speed and inhibitory control and differences in performance between true belief—positive desire versus false belief—negative desire in the first-order domain (Apperly et al., 2011).

Lagattuta et al. (2016) conducted a study involving children aged 4 to 10 years to investigate how children reason about mental states, specifically the interactions in their representations among thoughts, emotions, and decisions, and the role of executive functioning in these reasoning. The results showed that children between 3 and 7 years have an increasing tendency to explain the causes of decisions in terms of what people think (Lagattuta and Wellman, 2001). Lagattuta et al. (2016) also found that children aged 8–10 show greater valence alignment of thoughts, emotions, and decisions compared to younger participants. This means that older children exhibit greater consistency between positive thoughts, emotions, and decisions and between negative thoughts, emotions, and avoidant decisions. However, executive functions such as working memory and inhibitory control may also influence the interaction between thoughts, emotions, and decisions and may be involved in maintaining valence alignment (Lagattuta et al., 2016).

Interesting evidence on the development and relationship between ToM, cognitive and communicative skills has also been found in studies of the domain of lying (e.g., Cheung et al., 2015). Specifically, in preschoolers, first-order ToM was associated with self-motivated lying but not with other-motivated lying, which requires greater cognitive effort not only to inhibit the truth but also to consider the other person's interest (Talwar et al., 2017). Moreover, in a sample of 3- to 8-year-old children, only first-order ToM, but not second-order ToM, played a role in sincere and deceptive communicative acts (Bosco and Gabbatore, 2017).

Language and executive functions also significantly contribute to the development of Advanced ToM (Filippova and Astington, 2008; Lecce et al., 2017; Wilson et al., 2018). Osterhaus et al. (2016) highlighted the relationship between language, inhibitory abilities, and social reasoning and ambiguous reasoning in children aged 8 to 10 years (Osterhaus et al., 2016).

1.4 The present study

The objective of this study is to improve the understanding of the development of second-order reasoning from preschool when first-order reasoning is typically mastered (Miller, 2012), to early primary school, a school age in which advanced ToM forms start to emerge (Bianco et al., 2021; Osterhaus and Koerber, 2021b). In this way, we can track the developmental steps between first-order- and second-order-mastering. The specific purpose of the current study is to examine the ability to reason about mental states involving positive vs. negative desire and true vs. false belief in a second-order recursive thinking scenario. The focus on desires and beliefs is supported by evidence on first-order ToM, which placed these mental states at the basis of subsequent development (Apperly et al., 2011; Wellman and Liu, 2004). Our first hypothesis was that mastery of different types of traditional ToM tasks and of new ToM tasks at different ages could vary also depending on the level of reasoning explicitness required, the difficulty of the stories' structure and language, and the alignment between reality and the mental states of the characters (Beaudoin et al., 2020; Coull et al., 2006; Miller, 2022; Sullivan et al., 1994). To achieve this aim, traditional tasks (Perner and Wimmer, 1985; Sullivan et al., 1994; Castelli et al., 2000) were used as a point of comparison with second-order recursive stories constructed specifically for the present study.

Secondly, we hypothesized that second-order reasoning follows the same developmental pattern found for first-order reasoning, with the understanding of positive desire being achieved earlier than the understanding of negative desire, and with the mastering of true belief preceding false belief (Apperly et al., 2011; Wellman and Liu, 2004).

Finally, because the development of ToM interacts with executive functioning (Doenyas et al., 2018; Traverso et al., 2022; Wilson et al., 2018) and verbal abilities (Astington and Baird, 2012; Belacchi, 2022; Milligan et al., 2007), our third hypothesis concerns the presence of some associations between ToM development verbal abilities and executive functioning.

2 Materials and methods

2.1 Procedure

Participants were recruited from public schools in the North of Italy. The study was approved by the Ethical Committee for Research of the University of Bergamo (Report No. 1/2023 of 18th January 2023), and all ethical guidelines were followed (Associazione Italiana di Psicologia, 2022; APA, 2017; World Medical Association Declaration of Helsinki, 2013). Informed written consent was required for participation, and the document was provided to parents by teachers at school. All participants were allowed to withdraw at any time and were provided with the researchers' contact information for any questions or additional information. The inclusion criteria for the study required fluency in the Italian language and the absence of any neurodevelopmental disorders or developmental delays as reported by the parents. The study was conducted at school in three individual sessions. In the first session, children completed: a traditional second-order ToM task, 3 stories from the Belief × Desire II-order task, and a working memory task. In the second session, children completed a verbal ability task, another traditional second-order ToM task, and 3 more stories from the Belief × Desire II-order task. The third session consisted of an inhibition task, 2 stories from the Belief × Desire II-order task, and the Triangle task.

2.2 Participants

The study involved 59 children, 36 of whom were male, with age ranging from 5 to 8 years ($M = 6.52$, $SD = 0.79$). Group 1 consisted of 24 preschoolers in their last year of pre-primary education (aged 5;5 to 6;4), Group 2 consisted of 15 pupils in Grade 1 of primary school (aged 5;11 to 6;10), and Group 3 consisted of 20 pupils in Grade 2 of primary school (aged 6;11 to 7;10).

2.3 Measures

2.3.1 Verbal ability

Verbal skills were assessed using the Verbal Meaning (VM) subtests of the Primary Mental Ability (PMA) battery (Thurstone and Thurstone, 1965), composed of 32 items. Participants selected which of four pictures had the same meaning as a target word spoken aloud by the researcher. One point was given for each correct answer (range 0–32).

2.3.2 Executive functions

Executive functioning was assessed by testing inhibitory control and working memory. The Fruit Stroop task (Archibald and Kerns, 1999) consisted of three familiarization trials in which children were asked to name the four colors of rectangles, fruits, and vegetables (on both colored and uncolored stimulus pages). The fourth stimulus page included inhibitory control trials. Fruits and vegetable were presented with incorrect colors, such as a purple apple, and participants were asked to correctly name the color that each stimulus should have been. On each trial, children were asked to name the stimuli as quickly as possible within a time limit of 45 s. Scores were calculated by giving 1 point for each color correctly named within the time limit.

The study assessed working memory skills using a backward word recall task (Lanfranchi et al., 2004). Participants were asked to repeat a series of two to six words in reverse order. Each difficulty level had two trials, and 1 point was awarded for each correct backward recall (range 0–10).

2.3.3 Traditional ToM tasks

This study employed two classical second-order false belief tasks: the "*Ice cream seller*" task (Perner and Wimmer, 1985) and the *"Chocolate bar"* task (Sullivan et al., 1994). Each task consisted of a story that children could also follow along with some vignettes while the researcher read them. At the end of each story, children were asked a series of questions, including control and first-order questions, a second-order test question, and a justification question. Participants received 1 point for correctly answering the second-order test question and an additional point for providing the correct justification. The second-order question required a correct answer to the first-order question and at least two control questions. The range of second-order scores for each task was 0–2. To obtain a general score for the traditional second-order tasks, the scores on single tasks were summed (range 0–4). To allow statistical comparisons with success at the chance level, all scores were converted to a proportion of success ranging from 0 to 1.

2.3.4 Triangle task

In this ToM task (Castelli et al., 2000; White et al., 2011), participants were asked to describe what they thought was happening in a silent video clip in which a big red triangle and a small blue triangle made some movements. Children viewed 3 video clips in random order, each of which elicited mental state attributions through animations. Verbal descriptions were recorded and coded, with intentionality scores ranging from 0 (absence of mental state references) to 5 (elaborate reference to mental states). Two independent raters coded 25% of verbal descriptions, resulting in a calculated Cohen's kappa agreement of $k = 0.82$. The intentionality score ranges from 0 to 15.

2.3.5 Belief × Desire II-order task

In this new task developed appropriately for the current study, the researcher presented a set of 8 stories in a randomized order to investigate the development of second-order reasoning about beliefs and desires. The stories were constructed following what was done in first-order ToM (Apperly et al., 2011). The study manipulated 4 conditions: B + D+, B–D–, B + D–, B-D+. The acronym "D+" indicates positive desire and "D–" indicates negative desire; the acronym "B+" denotes true beliefs, and the acronym "B–" denotes false beliefs. Two stories were shown in each condition. Children could follow the stories on some vignettes while the researcher read them, and there were no time limits for answering the questions.

All tasks include two control questions. The first control question on a crucial plot of the story ensures that the difficulty of understanding the storyline did not affect performance. The second control question verifies whether participants correctly understood the desire of the characters in second-order reasoning. All the control questions were considered a prerequisite for the second-order question, and we scored 1 point if, besides control questions, the child also answered the second-order questions correctly (range 0–1 for each story). Scores were calculated for each condition B + D+, B + D–, B-D+, B–D– (range 0–2). To statistically compare scores with the chance level performance, scores were converted to a proportion of success for each condition (range 0–1). Items example can be found in Appendix A.

2.4 Data analysis

Data analysis was performed using Jamovi Software version 1.6.23. One-way ANOVAs were used to compare task scores across school grades. The one-sample *t*-test was used to compare children's performances to the chance level (0.50). Pearson and Spearman's correlations were used to identify potential associations between scores on different tasks, particularly with regard to the Belief × Desire II-order task in relation to traditional tasks and scores on verbal ability and executive functions. Repeated measures ANOVA was used to identify any significant interactions between the key factors of manipulation in the Belief × Desire II-order task.

3 Results

First, we analyzed the properties of the distribution of the scores displayed by the participants in the linguistic and executive functions tasks and in the traditional ToM tasks. As it can be seen in Table 1 the results indicated a negative skewness of verbal ability, suggesting that the majority of children in our sample achieved high scores on this task. The parameters of the other measurements presented in Table 1 suggested a distribution of data that does not significantly violate normality, as values of skewness and kurtosis between −1 and + 1 are considered acceptable. According to our first hypothesis concerning the possible role of tasks features in influencing the performance, Table 2 offers a first term of comparison between the distributions of scores on traditional tasks and the new one introduced in this study. The distribution of the scores on the Belief × Desire II-order tasks showed a positive skewness (>1), indicating a prevalence of lower scores in the sample for tasks that investigate false belief along with both positive and negative desire, as well as for true belief tasks that involve negative desire. As shown in Table 3, which illustrates the correlations between the Belief × Desire II-order tasks and traditional ToM measures, a positive correlation existed between the scores on the ToM triangles task and the B + D+ tasks. This correlation remained significant when working memory was included as a control variable, but not when inhibition acted as the control variable, *rho* = 0.174, $p = 0.191$. The One-way ANOVA detected some school grades differences in ToM reasoning assessed with the Triangle task, $F(2, 56) = 3.17$, $p = 0.049$, $\eta^2_{partial} = 0.102$. Group 3 ($M = 7.7$, $SD = 1.84$) scored higher, $p = 0.044$, CI [−1.39, −0.14], than Group 1 ($M = 5.75$, $SD = 2.83$). A similar significant difference related to school level concerned B + D+, $F(2, 56) = 5.80$, $p = 0.005$, $\eta^2_{partial} = 0.172$, where Group 3 ($M = 1.30$, $SD = 0.57$) performed better, $p = 0.006$, CI [−1.62, −0.35], than Group 1 ($M = 0.58$, $SD = 0.78$). Notably, there were no significant differences in school grades in the other Belief × Desire II-order task conditions B + D− ($p_s \geq 0.929$), B–D+ ($p_s \geq 0.311$), B–D− ($p_s \geq 0.230$), and on the traditional measures of second-order false belief, $p_s \geq 0.183$. Significant results that contribute to verify our first hypothesis are displayed also in Table 4 that illustrates the significantly below chance performance of all children (Younger and Older) on the traditional second-order tasks. However, when the second-order traditional stories were considered separated, only performance on the *Ice cream seller* story (Perner and Wimmer, 1985) was below chance level for both younger and older children. In the *Chocolate Bar* story (Sullivan et al., 1994), older children's scores were not below chance, but they were not above it either ($p = 0.66$). The results of the comparison of the two (younger and older) groups with the chance level on various types of Belief × Desire II-order tasks are presented in Table 5. Overall, both age groups of children scored below chance level on negative desire reasoning and/or false belief tasks. However, the older children performed above chance on second-order reasoning about true belief combined with positive desire, whereas the younger group performed below chance level on this type of task.

The age differences observed in Belief × Desire II-order tasks provide initial evidence that supports our second hypothesis, namely a second-order developmental pattern similar to those detected in studies concerning first-order reasoning. To deepen the significance of these results a repeated measures ANOVA was conducted with desire and belief as within-subject factors. There was a significant effect of desire, $F(1, 58) = 61.1$, $p < 0.001$, $\eta^2_{partial} = 0.513$. The Bonferroni *post hoc* comparisons revealed a significant ($p < 0.001$) mean difference (0.250) between positive and negative desire, favoring positive. The analysis also revealed a significant effect of belief, $F(1, 58) = 50.2$, $p < 0.001$, $\eta^2_{partial} = 0.464$, with Bonferroni post hoc comparisons indicating a significant ($p < 0.001$) mean difference (0.216) between true and false beliefs, favoring the true. The interaction belief × desire was also significant, $F(1, 58) = 24$, $p < 0.001$, $\eta^2_{partial} = 0.292$, and Bonferroni post hoc

TABLE 1 Descriptive statistics of verbal ability, executive functions and traditional ToM measures.

	Min	Max	*M*	SD	sk	ku
Verbal ability	5	30	23.3	4.70	−1.73	4.07
Inhibitory control	0	40	22.2	9.17	−0.79	0.28
Working memory	0	6	2.76	1.76	−0.07	−0.86
Second-order false belief	0	3	0.80	0.92	0.69	−0.84
Intentionality ToM triangle	0	12	6.64	2.65	−0.1	−0.02

TABLE 2 Descriptive statistics of Belief × Desire II-order task.

Types of ToM task	Min	Max	*M*	SD	sk	ku
B + D+	0	2	0.97	0.79	0.06	−1.37
B + D−	0	1	0.17	0.38	1.81	1.31
B–D+	0	1	0.24	0.43	1.27	−0.41
B–D−	0	1	0.03	0.18	5.29	26.9

TABLE 3 Spearman correlations between traditional ToM tasks and Belief × Desire II-order task.

Traditional measures	B + D+		B + D−		B−D+		B−D−	
	Rho	*p*	*Rho*	*p*	*Rho*	*p*	*Rho*	*p*
Second-order false belief	0.22	0.09	−0.03	0.82	−0.12	0.37	0.17	0.21
Intentionality ToM triangles	0.28	0.04	0.18	0.17	0.05	0.70	−0.12	0.36

TABLE 4 One-sample *t*-test performances below chance on traditional ToM tasks in younger and older children.

	N	*M*	*SD*	Student's *t*	*df*	*p*	Effect size Cohen's *d*
Second-order false belief (general score)							
Younger	36	0.17	0.22	−8.81	35	<0.001 [a]	−1.469
Older	23	0.24	0.24	−5.12	22	<0.001 [a]	−1.069
		M	*SD*	Student's *t*	df	*p*	Effect size Cohen's *d*
Younger	36						
Ice cream seller		0.11	0.24	−9.63	35	<0.001[a]	−0.761
Chocolate bar		0.24	0.33	−4.84	35	<0.001[a]	−0.807
		M	*SD*	Student's t	*df*	*p*	Effect Size Cohen's *d*
Older	23						
Ice cream seller		0.02	0.10	−22.0	22	<0.001[a]	−4.587
Chocolate bar		0.46	0.50	−0.42	22	0.34	−0.087

Younger: children under the age of 6;6. Older: children over the age of 6;7. [a]Population mean < 0.5.

TABLE 5 One sample *t*-test performances above and below chance in younger and older children.

Younger					Below chance	Above chance	
Second-order conditions	*M*	*SD*	Student's *t*	*df*	*p*	*p*	Effect size Cohen's *d*
B + D+	0.38	0.42	−1.78	35	0.042	0.958	−0.297
B + D−	0.08	0.19	−13.23	35	< 0.001	1.00	−2.205
B−D+	0.10	0.20	−12.04	35	< 0.001	1.00	−2.007
B−D−	0.03	0.12	−24.39	35	< 0.001	1.00	−4.065

Older					Below chance	Above chance	
Second-order conditions	*M*	*SD*	Student's *t*	*df*	*p*	*p*	Effect size Cohen's *d*
B + D+	0.65	0.28	2.61	22	0.992	0.008	0.545
B + D−	0.09	0.19	−10.2	22	< 0.001	1.000	−2.131
B−D+	0.15	0.24	−7.09	22	< 0.001	1.000	−1.478
B−D−	0.00	0.00	−Inf	22	< 0.001	1.000	−Inf

Younger: children under the age of 6;6. Older: children over the age of 6;7.

comparisons revealed significant differences between B + D+ and all other conditions (B + D−, $p < 0.001$, mean difference = 0.399; B−D+, $p < 0.001$, mean difference = 0.36; B−D−, $p < 0.001$, mean difference = 0.47), and between B−D+ and B−D− ($p = 0.012$, mean difference = 0.10). Not significant differences were found between B + D− and B−D+ ($p = 1.000$, mean difference = 0.03) and between B + D− and B−D− ($p = 0.118$, mean difference = 0.07). When the between-subject factor of two age groups was introduced, the

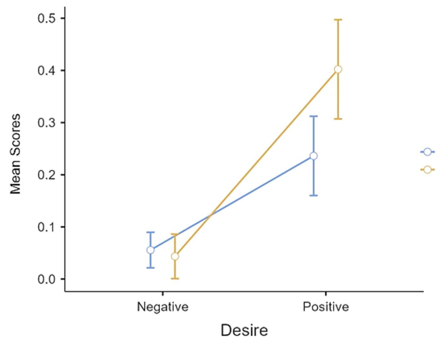


FIGURE 1
Interaction desire × age group.

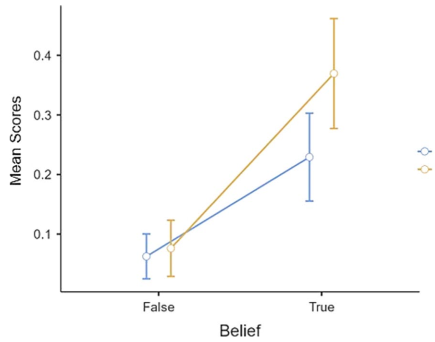


FIGURE 2
Interaction belief × age group.

results indicated a significant interaction desire × age group, $F(1, 57) = 8.31$, $p = 0.006$, $\eta^2_{partial} = 0.127$ (Figure 1). *Post hoc* comparisons on D+ comparing younger and older children were marginally significant ($p = 0.051$) with a mean difference of 0.17, suggesting that older children performed better than younger ones, although the results did not reach significant threshold for significance. Furthermore, post hoc comparisons revealed a significant ($p < 0.001$) mean difference between positive and negative desires in both younger and older groups, favoring positive desires. There was also a significant interaction between belief and age group, $F(1, 57) = 4.35$, $p = 0.041$, $\eta^2_{partial} = 0.071$ (Figure 2). Post hoc comparisons revealed a significant ($p < 0.001$) mean difference between true and false belief in both younger and older group comparisons, in favor of true belief.

As shown in Table 6, to verify our third hypothesis we analyzed possible associations between traditional ToM tasks and measures of executive functions. The Pearson correlations indicated a positive correlation between traditional second-order ToM tasks and both working memory and inhibition control. Additionally, the inhibition control task showed a positive correlation with the Triangle task. There was also an internal positive correlation between the two measures of executive functioning. Table 7 displays correlations between language/executive measures and the Belief × Desire II-order task. Specifically, the B + D+ tasks score was positively correlated with the inhibition score. Concerning this third hypothesis a one-way ANOVA revealed also significant differences in executive function scores between children of different school grades. Specifically, significant differences were found in scores on working memory, $F(2, 56) = 3.61$, $p = 0.033$, $\eta^2_{partial} = 0.114$, and inhibition tasks, $F(2, 56) = 16.0$, $p = <0.001$, $\eta^2_{partial} = 0.364$. *Post hoc* comparisons showed a significant mean difference, $p = 0.028$, 95% CI [−1.44, −0.19], in working memory between Group 1 ($M = 2.13$, $SD = 1.62$) and Group 3 ($M = 3.5$, $SD = 1.32$). For the inhibition task, there were significant differences, $p < 0.001$, between Group 1 ($M = 15.58$, $SD = 7.76$) and both Group 2 ($M = 26.8$, $SD = 5.13$), CI [−2.23, −0.79], and Group 3 ($M = 26.7$, $SD = 8.42$), CI [−2.16, −0.83].

4 Discussion

This study explores the attainment of second-order reasoning (Perner and Wimmer, 1985), which has been described in literature as an early stage in the development of advanced ToM (Osterhaus et al., 2016; Osterhaus and Koerber, 2021a). In this study, we investigate the potential interconnections between the understanding of different mental states (i.e., positive vs. negative desires, and true vs. false beliefs) and whether they are understood at different ages in interaction with executive functions and linguistic abilities. Moreover, the results provided insights into the potential influence of ToM task characteristics on the detection of second-order development in middle childhood.

As expected, there were differences in the performance of younger and older children on the traditional tasks, which are included in line with our first aim to provide a valuable comparison of new tasks' results. In the analysis of the individual stories, it was observed that the sample performance on the *Ice cream seller* story (Perner and Wimmer, 1985) was below chance level. However, the performance on the C*hocolate bar* story (Sullivan et al., 1994) was not below chance for the older group, indicating a possible lower level of difficulty for this task (Coull et al., 2006; Miller, 2022). For *The ice cream seller* story (Perner and Wimmer, 1985), the literature suggests that it is typically passed at the age of 7. Therefore, it is not surprising that the younger group, consisting of children aged 5–6.5 years, obtained low scores, and so did the older group, consisting of children younger than 7 years. Furthermore, while traditional second-order false belief tasks are typically considered mastered by age 7 (Hughes and Devine, 2015; Miller, 2009), some studies have shown that it is not until age 11 that all typically developing children are able to successfully complete second-order false belief task (Arslan et al., 2013). Additionally, *The ice cream seller* task has been found to be particularly challenging (Coull et al., 2006), even for children older than those in the present study (Braüner et al., 2020). Furthermore, Massaro et al. (2014) found no age effect on the performance of second-order false belief tasks in children aged 7, 8, and 11 years.

The ToM Triangle task showed a significant difference between the three school grades. In this task, children are asked to assign mental meaning to situations without making comparisons with reality, as highlighted in a recent paper (Lombardi et al., 2022). According to the results, Group 3 demonstrated a significantly higher level of achievement in this type of ToM ability compared to Group 1. This suggests that achievement of this ability is comparatively easier than second-order false belief reasoning. It is possible that the Triangle task is more effective in showing the improvement of ToM ability in this age group, while second-order false belief tasks may still be too challenging. It is noteworthy that some differences in second-order false belief performance are beginning to emerge. When comparing the false belief

TABLE 6 Pearson correlations between traditional ToM tasks and executive functions tasks.

Traditional measures	Second-order false belief		Intentionality ToM triangles		Working memory		Inhibitory control	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Second-order false belief	–	–						
Intentionality ToM triangles	0.03	0.84	–	–				
Working memory	0.26	0.05	0.08	0.54	–	–		
Inhibitory control	0.27	0.04	0.32	0.01	0.39	0.002	–	–

TABLE 7 Spearman correlations between verbal ability, executive functions and Belief × Desire II-order task.

Traditional measures	B + D+		B + D–		B–D+		B–D–	
	Rho	*p*	*Rho*	*p*	*Rho*	*p*	*Rho*	*p*
Verbal ability	0.31	0.017	0.19	0.14	0.14	0.29	0.003	0.98
Working Memory	0.18	0.16	−0.04	0.77	−0.05	0.73	−0.04	0.75
Inhibition	0.42	<0.001	0.05	0.69	0.16	0.22	−0.06	0.66

task with other types of task, such as the picture-sequencing task, it is important to consider that false belief tasks involve "*competing representations*" (Tsuji and Mitchell, 2019, p. 11) that require children to inhibit their own representation in order to succeed. On the contrary, in the picture-sequencing task, and even more so in the Triangle task, children are not asked to deal with a representation that competes with their own, making the cognitive demand lower and the task easier. However, further investigation is required because the literature suggests that the ages of 7–8 years is a sensitive period for the development of second-order reasoning (Bianco et al., 2021). We observed a positive correlation between the Triangle task and the second-order reasoning about true belief in a positive desire scenario where there is a concordance between reality and mental states. It could be hypothesized that the Triangle task and B + D+ tasks may require relatively less cognitive effort compared to the false belief tasks, which are known to have higher cognitive demands. The results suggest an increase in B + D+ second-order reasoning across the age range considered, as observed in the Triangle task.

To verify our second hypothesis, i.e., the existence of a similar pattern of development between first-order and second-order reasoning, we introduced a task that yielded interesting results. Performance on B + D+ tasks varied between age groups, with the oldest group of children performing above chance and the youngest group performing below chance. As mentioned above, in B + D+ tasks there are no "*competing representations*" (Tsuji and Mitchell, 2019, p. 11), or at least the representations are not opposed. However, further exploration is needed to investigate the role of inhibiting one's own desire in allowing the child to consider the character's desire (Rakoczy et al., 2007). When examining the results for B + D–, B–D+, and B–D–, it is apparent that the scores in the sample were skewed toward the lower end, indicating that most participants scored poorly, particularly in the B–D– condition. According to Friedman and Leslie (2005), first-order negative desire tasks were more challenging for 4-year-olds than traditional false belief tasks. In our sample, this condition was also found to be the most difficult. This may be due to the need for "*double inhibition*" (Friedman and Leslie, 2005, p. 222) to complete the task. However, it is possible that lower performance in B + D– tasks may be influenced by the difficulty of reasoning about avoiding something (D–). Moreover, the repeated measures ANOVA revealed a significant effect of desire, with children performing better on tasks involving positive than negative desires. This finding is consistent with previous studies on first-order thinking (Apperly et al., 2011), which also showed better performance on positive desires compared to negative ones. *Post hoc* comparisons for the significant effect desire × age group showed a marginally significant result for reasoning about positive desire, with older children performing better than younger. The repeated measures ANOVA on belief detected an effect of belief type. It was observed that true belief was better understood than false belief, which may replicate the developmental pattern detected in first-order reasoning, where the understanding of true/diverse belief precedes the understanding of false belief (Apperly et al., 2011; Peterson and Wellman, 2019; Rivas-Garcia et al., 2020; Wellman and Liu, 2004). Furthermore, results indicate that older children performed better in understanding true beliefs than false beliefs, and a similar pattern was found for younger children. According to Lagattuta et al. (2016), older children between the ages of 8 and 10 exhibited superior performance in aligning different mental states compared to their younger counterparts. Furthermore, Apperly et al. (2011) observed that 6–7- years old children still have difficulty and make errors in the negative conditions related to belief and desires in first-order belief-desire reasoning, but even adults struggle to perform optimally in this condition, as revealed by reaction times. It is possible to hypothesize that the combination of beliefs and desires in the same reasoning may be challenging for the children in our sample, which includes children younger than 8 years. This may be particularly true when dealing with negative desires and false beliefs, which are likely to require a higher cognitive demand and complex inhibitory processes.

The third aim was to explore the interactions between ToM executive functions and verbal ability. The majority of children scored high on the verbal ability task, preventing us from adequately exploring the role of verbal abilities in the Belief × Desire II-order task. It is recommended that future research employ measures that are more sensitive to individual differences in verbal ability at this age.

Regarding executive functions, our results are consistent with the existing literature (Bock et al., 2015). There was a significant difference in working memory and inhibition between Group 1 and 3. Furthermore, the inhibition score of Group 2 differed from that of Group 1 and slightly exceeded the mean of Group 3. The study suggests that performance on inhibition tasks was positively correlated with true belief tasks involving positive desire, but not with false belief and/or negative desire tasks that require the ability to inhibit the information about reality and their own desires (Apperly et al., 2011). However, there was a positive correlation between traditional second-order false belief tasks and executive functions. It is possible that the lack of correlation between inhibitory control and the combined false belief and desire tasks is due to the sample's overall difficulty with these tasks, as shown previously. The complexity of combined belief-desire reasoning may also be explained by studies of the interconnection between other cognitive abilities and second-order reasoning. The literature on first-order reasoning (Doenyas et al., 2018), might suggest that flexibility is more closely associated with diverse desires and beliefs than with inhibition. In their models of ToM in 7- and 8-year-old children, Im-Bolter et al. (2016) observed that mental attentional capacity played a significant role in addressing higher-order ToM reasoning. Future studies could potentially benefit from the inclusion of other executive functions measures, such as flexibility (Tsuji and Mitchell, 2019).

4.1 Limitations and future directions for research

The first limitation of the study is that the Belief × Desire II-order task was utilized for the first time. Therefore, further in-depth research is necessary to confirm its validity. To avoid any potential bias, the story contents were varied, but future studies should investigate whether certain task features may have an impact on scores (e.g., the length of the stories, the number of characters and scenes involved). As previously noted, a second limitation concerns the ceiling effect in the verbal ability measure. Future studies should employ a more sensitive measure and explore the associations between second-order ToM and different domains of linguistic and communicative abilities. Additionally, a larger sample size is required, also including children between the ages of 8 and 10 to gain a more comprehensive understanding of ToM development during middle childhood. This approach could provide valuable insights, especially considering the challenges that the current sample seems to face with second-order reasoning, which may be due to their young age. The final goal for future research could be the creation of a new measurement scale that can capture individual differences in the developmental trajectories of this fundamental and multi-componential ability, and the design of training and educational interventions to support ToM development.

5 Conclusion

This study aimed to address the lack of research on the development of ToM between first-order reasoning and more advanced ToM abilities. The main finding is that children are better at managing positive desires than negative ones in second-order scenarios and that they tend to understand true beliefs more easily than false beliefs, even in second-order reasoning. Our findings lay the groundwork for future research on the development of second-order reasoning, particularly in relation to different mental states (i.e., desires and beliefs) and their interactions with other developmental processes, such as executive functions. The ability to understand the reasoning behind others' desires and beliefs is a key component of ToM, which is fundamental in everyday interactions (Castelli et al., 2022). By deepening our understanding of ToM developmental trajectory, we can further explore how the components here investigated influence the quality of interpersonal relationships and support the development of emotional and social skills (Bianco and Castelli, 2023; Lecce and Devine, 2021). Additionally, a more nuanced understanding of ToM development enables educators to better interpret and respond to children's behaviors (Bianco and Lecce, 2016; Lecce et al., 2022; Valle et al., 2022). This insight, indeed, empowers educators to intervene effectively, offering appropriate stimuli to children of different ages (Bianco and Castelli, 2023; Lombardi et al., 2022). We also think that researchers starting from our study can contribute to this process by further investigating developmental ToM mechanisms, which will facilitate the creation of more targeted and effective intervention programs.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://doi.org/10.6084/m9.figshare.24593013.

Ethics statement

The studies involving humans were approved by Ethics Committee of University of Bergamo (Report No. 1/2023). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

FB: Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Writing – original draft, Writing – review & editing. AC: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. DM: Conceptualization, Methodology, Supervision, Writing – review & editing. AM: Conceptualization, Methodology, Supervision, Writing – review & editing. IC: Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We are grateful to all participants, their families, teachers, and schools for their collaboration.

Conflict of interest

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1525368/full#supplementary-material

References

APA. (2017). Ethical principles of psychologist and code of conduct. Available at: https://www.apa.org/ethics/code (Accessed February 20, 2025).

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., and Todd, S. (2011). Developmental continuity in theory of mind: speed and accuracy of belief-desire reasoning in children and adults. *Child Dev.* 82, 1691–1703. doi: 10.1111/j.1467-8624.2011.01635.x

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

Arslan, B., Taatgen, N., and Verbrugge, R. (2013). "Modeling developmental transitions in reasoning about false beliefs of others" in Proceedings of the 12th international conference on cognitive modelling, 72.

Arslan, B., Taatgen, N. A., and Verbrugge, R. (2017). Five-year-olds' systematic errors in second-order false belief tasks are due to first-order theory of mind strategy selection: a computational modeling study. *Front. Psychol.* 8:275. doi: 10.3389/fpsyg.2017.00275

Arslan, B., Verbrugge, R., Taatgen, N., and Hollebrandse, B. (2020). Accelerating the development of second-order false belief reasoning: a training study with different feedback methods. *Child Dev.* 91, 249–270. doi: 10.1111/cdev.13186

Associazione Italiana di Psicologia. (2022). Codice Etico per La Ricerca in Psicologia. Available at: https://aipass.org/wp-content/uploads/2023/02/Codice-Etico_luglio-2022.pdf (Accessed February 20, 2025).

Astington, J. W., and Baird, J. A. (2012). Why language matters for theory of mind. New York: Oxford University Press.

Beaudoin, C., Leblanc, É., Gagner, C., and Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Front. Psychol.* 10:2905. doi: 10.3389/fpsyg.2019.02905

Belacchi, C. (2022). "Teoria della Mente e linguaggio Come condividere emozioni e pensieri con le parole," in *La competenza sociale nei bambini. Valutare e promuovere la Teoria della Mente in età prescolare e scolare*. eds. I. Castelli, D. Massaro and A. Marchetti (Bergamo: Edizioni Junior), 101–134.

Bianco, F., and Castelli, I. (2023). The promotion of mature theory of mind skills in educational settings: a mini-review. *Front. Psychol.* 14:1197328. doi: 10.3389/fpsyg.2023.1197328

Bianco, F., and Lecce, S. (2016). Translating child development research into practice: can teachers foster children's theory of mind in primary school? *Br. J. Educ. Psychol.* 86, 592–605. doi: 10.1111/bjep.12125

Bianco, F., Lecce, S., and Banerjee, R. (2016). Conversations about mental states and theory of mind development during middle childhood: a training study. *J. Exp. Child Psychol.* 149, 41–61. doi: 10.1016/j.jecp.2015.11.006

Bianco, F., Lombardi, E., Lecce, S., Marchetti, A., Massaro, D., Valle, A., et al. (2021). Supporting Children's second-order recursive thinking and advanced ToM abilities: a training study. *J. Cogn. Dev.* 22, 561–584. doi: 10.1080/15248372.2021.1901712

Bianco, F., Lombardi, E., Massaro, D., Castelli, I., Valle, A., Marchetti, A., et al. (2019). Enhancing advanced theory of mind skills in primary school: a training study with 7- to 8-year-olds. *Infant Child Dev.* 28, 1–16. doi: 10.1002/icd.2155

Bock, A. M., Gallaway, K. C., and Hund, A. M. (2015). Specifying links between executive functioning and theory of mind during middle childhood: cognitive flexibility predicts social understanding. *J. Cogn. Dev.* 16, 509–521. doi: 10.1080/15248372.2014.888350

Bosco, F. M., and Gabbatore, I. (2017). Sincere, deceitful, and ironic communicative acts and the role of the theory of mind in childhood. *Front. Psychol.* 8:21. doi: 10.3389/fpsyg.2017.00021

Braüner, T., Blackburn, P., and Polyanskaya, I. (2020). Being deceived: information asymmetry in second-order false belief tasks. *Topics Cogn. Sci.* 12, 504–534. doi: 10.1111/tops.12422

Castelli, F., Happé, F., Frith, U., and Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* 12, 314–325. doi: 10.1006/nimg.2000.0612

Castelli, I., Massaro, D., and Marchetti, A. (2022). La competenza sociale nei bambini. Valutare e promuovere la Teoria della Mente in età prescolare e scolare. Bergamo: Edizioni Junior.

Cheung, H., Siu, T.-S. C., and Chen, L. (2015). The roles of liar intention, lie content, and theory of mind in children's evaluation of lies. *J. Exp. Child Psychol.* 132, 1–13. doi: 10.1016/j.jecp.2014.12.002

Coull, G. J., Leekam, S. R., and Bennett, M. (2006). Simplifying second-order belief attribution: what facilitates children's performance on measures of conceptual understanding? *Soc. Dev.* 15, 260–275. doi: 10.1046/j.1467-9507.2006.00340.x

Doenyas, C., Yavuz, H. M., and Selcuk, B. (2018). Not just a sum of its parts: how tasks of the theory of mind scale relate to executive function across time. *J. Exp. Child Psychol.* 166, 485–501. doi: 10.1016/j.jecp.2017.09.014

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

Friedman, O., and Leslie, A. M. (2005). Processing demands in belief-desire reasoning: inhibition or general difficulty? *Dev. Sci.* 8, 218–225. doi: 10.1111/j.1467-7687.2005.00410.x

Harris, P. L., de Rosnay, M., and Pons, F. (2005). Language and Children's understanding of mental states. *Curr. Dir. Psychol. Sci.* 14, 69–73. doi: 10.1111/j.0963-7214.2005.00337.x

Hogrefe, G.-J., Wimmer, H., and Perner, J. (1986). Ignorance versus false belief: a developmental lag in attribution of epistemic states. *Child Dev.* 57:567. doi: 10.2307/1130337

Hughes, C. (2016). Theory of mind grows up: reflections on new research on theory of mind in middle childhood and adolescence. *J. Exp. Child Psychol.* 149, 1–5. doi: 10.1016/j.jecp.2016.01.017

Hughes, C., and Devine, R. T. (2015). "A social perspective on theory of mind," in *Handbook of child psychology and developmental science: Socioemotional processes (7th ed)*. eds. M. E. Lamb and R. M. Lerner (Hoboken, NJ: Wiley), 564–609.

Im-Bolter, N., Agostino, A., and Owens-Jaffray, K. (2016). Theory of mind in middle childhood and early adolescence: different from before? *J. Exp. Child Psychol.* 149, 98–115. doi: 10.1016/j.jecp.2015.12.006

Lagattuta, K. H., Elrod, N. M., and Kramer, H. J. (2016). How do thoughts, emotions, and decisions align? A new way to examine theory of mind during middle childhood and beyond. *J. Exp. Child Psychol.* 149, 116–133. doi: 10.1016/j.jecp.2016.01.013

Lagattuta, K. H., and Wellman, H. M. (2001). Thinking about the past: early knowledge about links between prior experience, thinking, and emotion. *Child Dev.* 72, 82–102. doi: 10.1111/1467-8624.00267

Lanfranchi, S., Cornoldi, C., and Vianello, R. (2004). Verbal and visuospatial working memory deficits in children with down syndrome. *Am. J. Ment. Retard.* 109, 456–466. doi: 10.1352/0895-8017(2004)109<456:VAVWMD>2.0.CO;2

Lecce, S., and Devine, R. (2021). Theory of mind at school: academic outcomes and the influence of the school context. *Infant Child Dev.* 31, 1–7. doi: 10.1002/icd.2274

Lecce, S., Bianco, F., Demicheli, P., and Cavallini, E. (2014). Training preschoolers on first-order false belief understanding: transfer on advanced ToM skills and Metamemory. *Child Dev.* 85, 2404–2418. doi: 10.1111/cdev.12267

Lecce, S., Bianco, F., Devine, R. T., and Hughes, C. (2017). Relations between theory of mind and executive function in middle childhood: a short-term longitudinal study. *J. Exp. Child Psychol.* 163, 69–86. doi: 10.1016/j.jecp.2017.06.011

Lecce, S., Ronchi, L., and Devine, R. T. (2022). Mind what teacher says: teachers' propensity for mental-state language and children's theory of mind in middle childhood. *Soc. Dev.* 31, 303–318. doi: 10.1111/sode.12552

Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.* 58, 259–289. doi: 10.1146/annurev.psych.58.110405.085654

Lockl, K., and Schneider, W. (2007). Knowledge about the mind: links between theory of mind and later Metamemory. *Child Dev.* 78, 148–167. doi: 10.1111/j.1467-8624.2007.00990.x

Lombardi, E., Valle, A., Bianco, F., Castelli, I., Massaro, D., and Marchetti, A. (2022). Supporting mentalizing in primary school children: the effects of thoughts in mind project for children (TiM-C) on metacognition, emotion regulation and theory of mind. *Cognit. Emot.* 36, 975–986. doi: 10.1080/02699931.2022.2067521

Massaro, D., Castelli, I., Sanvito, L., and Marchetti, A. (2014). The 'I knew it all along' phenomenon: second-order false belief understanding and the curse of knowledge in primary school children. *Eur. J. Psychol. Educ.* 29, 311–326. doi: 10.1007/s10212-013-0200-7

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychol. Bull.* 135, 749–773. doi: 10.1037/a0016854

Miller, S. A. (2022). *Advanced theory of mind*. New York: Oxford University Press.

Miller, S. A. (2012). *Theory of mind: Beyond the preschool years*. New York, NY: Psychology Press.

Milligan, K., Astington, J. W., and Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Dev.* 78, 622–646. doi: 10.1111/j.1467-8624.2007.01018.x

Osterhaus, C., and Bosacki, S. L. (2022). Looking for the lighthouse: a systematic review of advanced theory-of-mind tests beyond preschool. *Dev. Rev.* 64:101021. doi: 10.1016/j.dr.2022.101021

Osterhaus, C., and Koerber, S. (2021a). Social cognition during and after kindergarten: the relations between first-order and advanced theories of mind. *Eur. J. Dev. Psychol.* 18, 573–592. doi: 10.1080/17405629.2020.1820861

Osterhaus, C., and Koerber, S. (2021b). The development of advanced theory of mind in middle childhood: a longitudinal study from age 5 to 10 years. *Child Dev.* 92, 1872–1888. doi: 10.1111/cdev.13627

Osterhaus, C., Koerber, S., and Sodian, B. (2016). Scaling of advanced theory-of-mind tasks. *Child Dev.* 87, 1971–1991. doi: 10.1111/cdev.12566

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

Peterson, C. C., and Wellman, H. M. (2019). Longitudinal theory of mind (ToM) development from preschool to adolescence with and without ToM delay. *Child Dev.* 90, 1917–1934. doi: 10.1111/cdev.13064

Peterson, C. C., Wellman, H. M., and Slaughter, V. (2012). The mind behind the message: advancing theory-of-mind scales for typically developing children, and those with deafness, autism, or Asperger syndrome. *Child Dev.* 83, 469–485. doi: 10.1111/j.1467-8624.2011.01728.x

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/S0140525X00076512

Rakoczy, H., Warneken, F., and Tomasello, M. (2007). "This way!", "no! That way!"—3-year olds know that two people can have mutually incompatible desires. *Cogn. Dev.* 22, 47–68. doi: 10.1016/j.cogdev.2006.08.002

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

Rivas-Garcia, S., Bateman, A., and Caracuel, A. (2020). Validation of the ToMas-child scale for the assessment of the theory of mind in a Group of Spanish Speaking Children Aged 3 to 7 years from Spain. *Dev. Neuropsychol.* 45, 232–245. doi: 10.1080/87565641.2020.1764567

Sullivan, K., Zaitchik, D., and Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Dev. Psychol.* 30, 395–402. doi: 10.1037/0012-1649.30.3.395

Talwar, V., Crossman, A., and Wyman, J. (2017). The role of executive functioning and theory of mind in children's lies for another and for themselves. *Early Child. Res. Q.* 41, 126–135. doi: 10.1016/j.ecresq.2017.07.003

Thurstone, T. G., and Thurstone, L. L. (1965). P.M.A. primary mental abilities. *Organizzazioni Speciali*.

Traverso, L., Usai, M. C., Viterbori, P., and Zanobini, M. (2022). "Ti ho in mente: la relazione tra funzioni esecutive e Teoria della Mente," in *La competenza sociale nei bambini. Valutare e promuovere la Teoria della Mente in età prescolare e scolare*. eds. I. Castelli, D. Massaro and A. Marchetti (Bergamo: Edizioni Junior), 75–100.

Tsuji, H., and Mitchell, P. (2019). Modelling the executive components involved in processing false belief and mechanical/intentional sequences. *Br. J. Dev. Psychol.* 37, 184–198. doi: 10.1111/bjdp.12266

Valle, A., Rinaldi, T., Greco, A., Pianta, R., Castelli, I., and Marchetti, A. (2022). Mentalization and attachment in educational relationships at primary school. *Ricerche Psicol.* 1, 1–23. doi: 10.3280/rip2022oa13226

Wellman, H. M. (2012). Theory of mind: better methods, clearer findings, more development. *Eur. J. Dev. Psychol.* 9, 313–330. doi: 10.1080/17405629.2012.680297

Wellman, H. M., and Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Dev.* 75, 523–541. doi: 10.1111/j.1467-8624.2004.00691.x

White, S. J., Coniston, D., Rogers, R., and Frith, U. (2011). Developing the Frith-Happé animations: a quick and objective test of theory of mind for adults with autism. *Autism Res.* 4, 149–154. doi: 10.1002/aur.174

Wilson, J., Andrews, G., Hogan, C., Wang, S., and Shum, D. H. K. (2018). Executive function in middle childhood and the relationship with theory of mind. *Dev. Neuropsychol.* 43, 163–182. doi: 10.1080/87565641.2018.1440296

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128. doi: 10.1016/0010-0277(83)90004-5

World Medical Association Declaration of Helsinki (2013). *JAMA* 310:2191. doi: 10.1001/jama.2013.281053

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 19 March 2025
DOI 10.3389/fpsyg.2025.1534125

Check for updates

OPEN ACCESS

EDITED BY
Sara Isernia,
Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy

REVIEWED BY
Irene Ceccato,
G. d'Annunzio University of Chieti and Pescara, Italy
Federica Rossetto,
Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy

*CORRESPONDENCE
Susan A. J. Birch
✉ sbirch@psych.ubc.ca

RECEIVED 25 November 2024
ACCEPTED 25 February 2025
PUBLISHED 19 March 2025

CITATION
Birch SAJ, Stewardson CI, Rho K, Kataria A, Craig SM, Phan MDH, Savi I, Voronkova K, Lee J, Choudhary G and Torjani D (2025) Targeting cognitive biases to improve social cognition and social emotional health. *Front. Psychol.* 16:1534125. doi: 10.3389/fpsyg.2025.1534125

Targeting cognitive biases to improve social cognition and social emotional health

Susan A. J. Birch*, Charlotte I. Stewardson, Kate Rho, Ashali Kataria, Shannon M. Craig, Minh D. H. Phan, Irene Savi, Kseniia Voronkova, Jenny Lee, Gayatri Choudhary and Diba Torjani

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

KEYWORDS

social cognition, cognitive bias, theory of mind, social emotional health, cognitive debiasing, intervention, education, development

1 Introduction

Social cognitive skills are crucial for understanding and navigating human interactions, enabling us to process, interpret, and respond to social information (Arioli et al., 2018). A key component of these skills is *theory of mind*, which involves inferring and reasoning about one's own and others' mental states, including beliefs, intentions, desires, thoughts, and emotions (Premack and Woodruff, 1978; Wimmer and Perner, 1983). Theory of mind is essential in almost every social interaction, as it helps us understand human actions (e.g., Baron-Cohen, 1995; Frith and Frith, 2005), underpins cultural learning (e.g., Henrich, 2004; Herrmann et al., 2007), and is vital for effective communication and social decision-making (e.g., Baron-Cohen, 1995; Birch et al., 2017; Haddock and Birch, 2024). Theory of mind has also been shown to promote prosocial behavior (e.g., Imuta et al., 2016) and reduce prejudice (Shih et al., 2009).

Philosopher Elbert G. Hubbard aptly stated, "If men [sic] could only know each other, they would never idolize nor hate" (Hubbard, 1911, p. 13). We interpret this to mean that a rich understanding of each other's perspectives fosters greater social harmony and social emotional health. More specifically, we believe that by using *theory of mind* to understand others' perspectives, people can recognize their shared humanity and overcome the tendencies to either idealize or condemn others. This understanding may also help reduce biases and assumptions that lead to flawed social judgments, such as the "black-and-white" thinking in which others are seen as either flawless or completely flawed.

Consistent with Hubbard's sentiments, we propose that interventions enhancing social cognitive skills can significantly improve social-emotional health. Furthermore, we predict that the most successful interventions will incorporate strategies to minimize cognitive biases—systematic errors in thinking that affect decision-making and behavior (Tversky and Kahneman, 1974; for a review, see Ellis, 2018). We support this view by briefly reviewing research that shows: (a) enhancing social cognition improves various aspects of social-emotional health, (b) cognitive biases play a critical role in the link between social cognition and social-emotional health, and (c) strategies for reducing cognitive biases have tremendous promise for enhancing social cognition and social emotional health.

2 Improving social cognition can improve social-emotional health

Research has consistently demonstrated that theory of mind abilities are pivotal for developing and maintaining social relationships, particularly during childhood (Dunn and Cutting, 1999; Peterson et al., 2016; Etel and Slaughter, 2019; for reviews see Repacholi and Slaughter, 2003; Haddock and Birch, 2024). Evidence on the relationship between theory of mind and social competencies in adults, however, has been somewhat more mixed (Bora and Berk, 2016; Davis, 1983; Livingston et al., 2024; Wolgast et al., 2020). Generally speaking, individual differences in theory of mind are present early and continue into adulthood, with more advanced theory of mind predicting several positive outcomes (Dunn and Cutting, 1999; Repacholi and Slaughter, 2003). For example, higher theory of mind scores are associated with greater social understanding, higher levels of empathy, and more prosocial behavior, leading to reduced interpersonal conflicts and increased relationship satisfaction (e.g., Davis, 1983; Repacholi and Slaughter, 2003; for two meta-analyses see Imuta et al., 2016; Slaughter et al., 2015). Similarly, more advanced theory of mind has been associated with increased cooperation (Etel and Slaughter, 2019) as well as increased communication and reduced peer conflict (Dunn and Cutting, 1999; Haddock and Birch, 2024). Studies by Peterson et al. (2015) demonstrated that higher theory of mind is associated with increased self-esteem and higher quality friendships. Peterson et al. (2016) also found that children's theory of mind understanding independently predicted social skills above and beyond age, gender, and verbal ability. Furthermore, more advanced theory of mind has also been shown to reduce the risk of social adversity, such as bullying and social exclusion (Bosacki et al., 2020; Smith, 2017). More advanced theory of mind also appears to act as a protective factor against trauma and adversity (e.g., Cadamuro et al., 2016; Hughes and Ensor, 2006, 2007). Conversely, poor theory of mind skills are associated with greater psychological distress (Wolgast et al., 2020), more emotional symptoms, and increased loneliness (Caputi and Schoenborn, 2018). This latter result is especially noteworthy given longitudinal studies linking loneliness to a variety of negative health outcomes, including poorer sleep quality (Cacioppo et al., 2002), and increased depressive symptoms (Cacioppo et al., 2010). Even in adulthood, theory of mind predicts emotional symptoms such as sadness and depression. A meta-analysis of 18 studies examining the relationship between theory of mind and Major Depressive Disorder in adults revealed that deficits in theory of mind can be a risk factor for depression and accompanying psychosocial impairment, with the level of theory of mind impairment predicting symptom severity (Bora and Berk, 2016).

3 Maximize social cognition by minimizing cognitive biases

Decades of research across the psychological sciences have shown that cognitive biases play a critical role in shaping our perceptions, decisions, and interactions, influencing nearly every aspect of human interaction (Tversky and Kahneman, 1974; Kahneman, 2011; for a review, see Ellis, 2018). These cognitive biases are normal by-products of how the mind works; nonetheless, individual differences in the magnitude of these biases predict a range of outcomes. Cognitive biases lead to errors in decision-making and social judgments, impede communication, contribute to maladaptive behaviors, and even play a role in mental health conditions like depression (Beck, 1979; Kahneman, 2011; Nisbett and Ross, 1980; Tversky and Kahneman, 1974).

Not surprisingly, the way we think *about others* and their mental states is also vulnerable to cognitive biases. Given their social elements, cognitive biases are sometimes referred to as social cognitive biases. Social cognitive biases, systematic tendencies, or errors, in the way we think about others and their mental states, can be particularly damaging to interpersonal relationships, impair communication, and lead to poor social decision-making (e.g., Birch and Bernstein, 2007; Nickerson, 1999; Savitsky et al., 2011). For instance, consider the spotlight effect which occurs when individuals overestimate the extent to which others notice and evaluate their actions and appearance (Gilovich et al., 2000). This can lead to heightened self-consciousness and increased social anxiety, as individuals mistakenly believe they are under scrutiny. For example, in contexts like volleyball games and video games, participants overestimated how much their teammates notice differences in their performance compared to a typical game and anticipated harsher evaluations than were actually given (Gilovich et al., 2000, 2002). This tendency for individuals to feel that they are the center of attention, especially in potentially unfavorable situations, is linked to increased self-consciousness and social anxiety (e.g., Brown and Stopa, 2007).

Another cognitive bias that plays a clear role in social cognition is the curse of knowledge bias. The curse of knowledge bias refers to the tendency to be swayed by one's knowledge when attempting to reason about a more naive perspective (e.g., Birch and Bloom, 2003; Bernstein et al., 2004; Camerer et al., 1989; Fischhoff, 1977; Taylor et al., 1994; Sutherland and Cimpian, 2015; for a meta-analyses of 122 studies see Christensen-Szalanski and Willham, 1991). A classic example of the curse of knowledge bias (sometimes called 'hindsight bias') is when adults who know the outcome of an event (e.g., a sports game, an election, or a battle) overestimate how likely others are to predict that outcome. In contrast, adults who do not know the event's outcome tend to make more accurate estimates of what others will predict (e.g., Blank et al., 2003; Fischhoff, 1975; Fischhoff and Beyth, 1975, for review see Ghrear et al., 2016; for a meta-analyses see Guilbault et al., 2004). Given that the curse of knowledge bias leads individuals to overestimate how common their knowledge is, it regularly impacts communication and social judgments in various ways (e.g., Birch, 2005; Camerer et al., 1989). For example, experts often assume that their specialized knowledge is more widely understood than it is, which can hinder effective communication and lead to misunderstandings (Hinds, 1999). Importantly, research has shown that one of the most widely-used tasks to measure theory of mind, the classic 'false belief task,' is heavily influenced by the curse of knowledge bias. Although the curse of knowledge and false belief reasoning appear to be independent constructs with different developmental patterns (Bernstein, 2021), experimentally reducing this bias has been shown to improve false belief reasoning in both children (Ghrear et al., 2021) and adults (Birch and Bloom, 2007; see also Ghrear et al., 2020; Keysar et al., 2003). Importantly, interventions that provide contextual feedback about others' perspectives appear particularly effective at minimizing this type of 'egocentric bias' in adults (Damen et al., 2021).

Another well-documented cognitive bias in social cognition is the hostile attribution bias, which refers to the tendency to interpret ambiguous or neutral social cues as being intentionally hostile or aggressive (Dodge and Crick, 1990). The hostile attribution bias can significantly affect how individuals perceive and react to social interactions, because it shapes how people interpret the intentions of others. For instance, someone prone to this bias may interpret an accidental bump in a crowded hallway as a deliberate act of aggression and react aggressively in response. This bias overlaps with a broader phenomenon known as 'interpretation bias,' a widely studied bias in clinical research. Interpretation bias is a type of negativity bias involving the tendency to interpret ambiguous or neutral information in a negative manner. This bias is associated with higher levels of stress, anxiety, and depression (e.g., Mathews and MacLeod, 2002). In fact, research shows this cognitive bias is not merely associated with mood disorders but actively contributes to their development and maintenance (Gotlib and Joormann, 2010; Mathews and Mackintosh, 2000; Mathews and MacLeod, 2002; Kindt and Van Den Hout, 2001). Interpretation biases may also be associated with psychotic symptoms. That is, negative interpretation bias such as hostile attribution bias tends to be more pronounced among individuals who are experiencing both clinical and subclinical levels of psychosis, though the quality of some of these studies varies (for a review, see Samson et al., 2024; see also Beck and Clark, 1997). Interestingly, it has been theorized that negative interpretation biases might explain the increased rates of social withdrawal among individuals with subclinical and clinical levels of psychotic symptoms (e.g., negative interpretations of social interactions could reinforce a tendency to isolate; Rector and Beck, 2002). This latter observation reinforces the point that minimizing social cognitive biases have tremendous potential for improving social emotional health. The aforementioned biases are only a few examples of cognitive biases that affect social cognition. There are many others; several of which may share underlying mechanisms (e.g., Birch and Bloom, 2003; Birch and Bernstein, 2007; Oeberst and Imhoff, 2023; Tversky and Kahneman, 1974).

Fortunately, the ability to minimize many cognitive biases has been well-documented (e.g., Ross et al., 1977; Hirsch et al., 2018; Hooper et al., 2015; Macrae et al., 2016). For example, training to reduce negativity biases, such as interpretation bias, in individuals with a history of depression, has been shown to lessen the severity of depressive symptoms (Hirsch et al., 2018; Hofmann et al., 2012). Similarly, cognitive debiasing interventions with individuals with hostile attribution bias have also been shown to be effective (e.g., Hiemstra et al., 2018). For individuals with schizophrenia spectrum disorders, metacognitive training (MCT), an evidence-based intervention addressing cognitive biases over 8 to 16 sessions, has been shown to effectively improve global social cognition and theory of mind, with adapted versions being used with other clinical populations such as individuals with major depressive disorder, obsessive-compulsive disorder and borderline personality disorder (for a review, see Hotte-Meunier et al., 2024). While some debiasing techniques involve lengthy and/or implicit debiasing techniques, approaches that educate individuals about cognitive biases and/or offer strategies to lessen them can also be highly effective (e.g., Morewedge et al., 2015; Gilovich et al., 2000; van Brussel et al., 2021). For instance, even a brief 30–60 min intervention educating individuals about biases and ways to address them resulted in significant bias reductions for at least 2 to 3 months (Morewedge et al., 2015). Similar research suggests that game-based formats and spaced reminders may be especially beneficial for minimizing bias (Clegg et al., 2014). These latter examples did not specifically examine the broader benefits for social cognition, nonetheless, we believe these types of debiasing strategies hold great promise for enhancing social cognition and several facets of social-emotional health (see Craig et al., 2024 for a recent review).

4 Conclusion

Reasoning about the minds of others is multifaceted—it is complex and nuanced. A recent review of theory of mind measures suggested that there are at least 39 different theory of mind sub-abilities (Beaudoin et al., 2020). Just as researchers should avoid relying on a single measure of theory of mind (e.g., Bloom and German, 2000; Haddock and Birch, 2024), we should also refrain from depending on any single intervention approach. Vast individual differences exist in people's strengths and limitations in reasoning about the minds of others. As such, we believe combining multiple strategies is the best way to address the multifaceted nature of theory of mind and the unique and diverse challenges individuals face.

Notably, some cognitive biases appear to play an even greater role during childhood and early adolescence than in adulthood (e.g., Birch, 2005; Bernstein et al., 2011; Ghrear et al., 2021, 2020), highlighting the importance of introducing debiasing strategies in younger populations. Educating parents and teachers about cognitive debiasing strategies also has considerable merit and can provide valuable indirect benefits in situations where directly teaching strategies to very young children might be challenging. For instance, Gehlbach and Vriesema (2019) suggest that educating individuals about cognitive biases and related theories equips them with tools to identify and create learning opportunities for children. These opportunities encourage children to reassess their perspectives during social interactions and conflicts, ultimately helping them reduce their biases. Addressing these biases early in development has the greatest potential to prevent social-emotional problems and yield the most long-term benefits—for individuals and for society as a whole.

In conclusion, it is our opinion that intervention approaches can maximize social cognition by minimizing cognitive biases. To be clear, we are not advocating a cognitive debiasing approach should replace existing intervention techniques, but rather that cognitive debiasing strategies be *integrated* with existing approaches. We believe that the most effective interventions for enhancing social cognition and social emotional health will combine existing methods with education on cognitive biases and concrete strategies to overcome them.

Author contributions

SB: Writing – original draft, Writing – review & editing. CS: Writing – review & editing. KR: Writing – review & editing. AK: Writing – review & editing. SC: Writing – review & editing. MP: Writing – review & editing. IS: Writing – review & editing. KV: Writing – review & editing. JL: Writing – review & editing. GC: Writing – review & editing. DT: Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Arioli, M., Crespi, C., and Canessa, N. (2018). Social cognition through the lens of cognitive and clinical neuroscience. *Biomed. Res. Int.* 2018:4283427. doi: 10.1155/2018/4283427

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Beaudoin, C., Leblanc, É., Gagner, C., and Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Front. Psychol.* 10:2905. doi: 10.3389/fpsyg.2019.02905

Beck, A. T. (1979). *Cognitive Therapy of Depression*. New York, NY: Guilford Press.

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

Bernstein, D. M. (2021). Hindsight bias and false-belief reasoning from preschool to old age. *Dev. Psychol.* 57, 1387–1402. doi: 10.1037/dev0001226

Bernstein, D. M., Atance, C., Loftus, G. R., and Meltzoff, A. (2004). We saw it all along: visual hindsight bias in children and adults. *Psychol. Sci.* 15, 264–267. doi: 10.1111/j.0963-7214.2004.00663.x

Bernstein, D. M., Erdfelder, E., Meltzoff, A. N., Peria, W., and Loftus, G. R. (2011). Hindsight bias from 3 to 95 years of age. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 378–391. doi: 10.1037/a0021971

Birch, S. A. J. (2005). When knowledge is a curse: children's and adults' reasoning about mental states. *Curr. Dir. Psychol. Sci.* 14, 25–29. doi: 10.1111/j.0963-7214.2005.00328.x

Birch, S. A. J., and Bernstein, D. M. (2007). What can children tell us about hindsight bias: a fundamental constraint on perspective-taking? *Soc. Cogn.* 25, 98–113. doi: 10.1521/soco.2007.25.1.98

Birch, S. A. J., and Bloom, P. (2003). Children are cursed: an asymmetric bias in mental-state attribution. *Psychol. Sci.* 14, 283–286. doi: 10.1111/1467-9280.03436

Birch, S. A. J., and Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychol. Sci.* 18, 382–386. doi: 10.1111/j.1467-9280.2007.01909.x

Birch, S. A. J., Li, V., Haddock, T., Ghrear, S. E., Brosseau-Liard, P., Baimel, A., et al. (2017). Perspectives on perspective taking: how children think about the minds of others. *Adv. Child Dev. Behav.* 52, 185–226. doi: 10.1016/bs.acdb.2016.10.005

Blank, H., Fischer, V., and Erdfelder, E. (2003). Hindsight bias in political elections. *Memory* 11, 491–504. doi: 10.1080/09658210244000513

Bloom, P., and German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77, B25–B31. doi: 10.1016/S0010-0277(00)00096-2

Bora, E., and Berk, M. (2016). Theory of mind in major depressive disorder: a meta-analysis. *J. Affect. Disord.* 191, 49–55. doi: 10.1016/j.jad.2015.11.023

Bosacki, S., Moreira, F. P., Sitnik, V., Andrews, K., and Talwar, V. (2020). Theory of mind, self-knowledge, and perceptions of loneliness in emerging adolescents. *J. Genet. Psychol.* 181, 14–31. doi: 10.1080/00221325.2019.1687418

Brown, M. A., and Stopa, L. (2007). The spotlight effect and the illusion of transparency in social anxiety. *J. Anxiety Disord.* 21, 804–819. doi: 10.1016/j.janxdis.2006.11.006

Cacioppo, J. T., Hawkley, L. C., Berntson, G. G., Ernst, J. M., Gibbs, A. C., and Stickgold, R. (2002). Lonely days invade the night: social modulation of sleep efficiency. *Psychol. Sci.* 13, 385–388. doi: 10.1111/1467-9280.00469

Cacioppo, J. T., Hawkley, L. C., and Thisted, R. A. (2010). Perceived social isolation makes me sad: five-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago health, aging, and social relations study. *Psychol. Aging* 25, 453–463. doi: 10.1037/a0017216

Cadamuro, A., Versari, A., Vezzali, L., and Trifiletti, E. (2016). Preventing the detrimental effect of posttraumatic stress in young children: the role of theory of mind in the aftermath of a natural disaster. *Eur. J. Dev. Psychol.* 13, 52–66. doi: 10.1080/17405629.2015.1055240

Camerer, C., Loewenstein, G., and Weber, M. (1989). The curse of knowledge in economic settings: an experimental analysis. *J. Polit. Econ.* 97, 1232–1254. doi: 10.1086/261651

Caputi, M., and Schoenborn, H. (2018). Theory of mind and internalizing symptoms during middle childhood and early adolescence: the mediating role of coping strategies. *Cogent Psychol.* 5, 1–15. doi: 10.1080/23311908.2018.1487270

Christensen-Szalanski, J. J., and Willham, C. F. (1991). The hindsight bias: a meta-analysis. *Organ. Behav. Hum. Decis. Process.* 48, 147–168. doi: 10.1016/0749-5978(91)90010-Q

Clegg, B. A., Martey, R. M., Stromer-Galley, J., Kenski, K., Saulnier, T., Folkestad, J. E., et al. (2014). "Game-based training to mitigate three forms of cognitive bias," in *Proceedings of Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)* (Orlando, FL), 1–12. Available online at: https://www.researchgate.net/publication/264158816_Game-based_Training_to_Mitigate_Three_Forms_of_Cognitive_Bias

Craig, S. M., Kataria, A., Rho, K., Voronkova, K., Phan, M. D. H., Lee, J., et al. (2024). Changing minds: reducing cognitive biases to enhance psychological health. *Psychol. J.* 6, 1–7. doi: 10.31038/PSYJ.2024654

Damen, D., van Amelsvoort, M., van der Wijst, P., Pollmann, M., and Krahmer, E. (2021). Lifting the curse of knowing: how feedback improves perspective-taking. *Q. J. Exp. Psychol.* 74, 1054–1069. doi: 10.1177/1747021820987080

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

Dodge, K. A., and Crick, N. R. (1990). Social information-processing bases of aggressive behavior in children. *Personal. Soc. Psychol. Bull.* 16, 8–22. doi: 10.1177/0146167290161002

Dunn, J., and Cutting, A. L. (1999). Understanding others, and individual differences in friendship interactions in young children. *Soc. Dev.* 8, 201–219. doi: 10.1111/1467-9507.00091

Ellis, G. (2018). "So, what are cognitive biases?" in: *Cognitive Biases in Visualizations*, ed. G. Ellis (Springer Nature, Switzerland), 1–12. doi: 10.1007/978-3-319-95831-6_1

Etel, E., and Slaughter, V. (2019). Theory of mind and peer cooperation in two play contexts. *J. Appl. Dev. Psychol.* 60, 87–95. doi: 10.1016/j.appdev.2018.11.004

Fischhoff, B. (1975). Hindsight is not equal to foresight: the effect of outcome knowledge on judgment under uncertainty. *J. Exp. Psychol.* 1, 288–299. doi: 10.1037//0096-1523.1.3.288

Fischhoff, B. (1977). Perceived informativeness of facts. *J. Exp. Psychol.* 3, 349–358. doi: 10.1037/0096-1523.3.2.349

Fischhoff, B., and Beyth, R. (1975). "I knew it would happen: "remembered probabilities of once–future things. *Organ. Behav. Hum. Perform.* 13, 1–16. doi: 10.1016/0030-5073(75)90002-1

Frith, C., and Frith, U. (2005). Theory of mind. *Curr. Biol.* 15, R644–R645. doi: 10.1016/j.cub.2005.08.041

Gehlbach, H., and Vriesema, C. C. (2019). Meta-bias: a practical theory of motivated thinking. *Educ. Psychol. Rev.* 31, 65–85. doi: 10.1007/s10648-018-9454-6

Ghrear, S., Baimel, A., Haddock, T., and Birch, S. A. J. (2021). Are the classic false belief tasks cursed? Young children are just as likely as older children to pass a false belief task when they are not required to overcome the curse of knowledge. *PLoS ONE* 16:e0244141. doi: 10.1371/journal.pone.0244141

Ghrear, S., Fung, K., Haddock, T., and Birch, S. A. J. (2020). Only familiar information is a "curse": children's ability to predict what their peers know. *Child Dev.* 92, 54–75. doi: 10.1111/cdev.13437

Ghrear, S. E., Birch, S. A., and Bernstein, D. M. (2016). Outcome knowledge and false belief. *Front. Psychol.* 7:118. doi: 10.3389/fpsyg.2016.00118

Gilovich, T., Kruger, J., and Medvec, V. H. (2002). The spotlight effect revisited: overestimating the manifest variability of our actions and appearance. *J. Exp. Soc. Psychol.* 38, 93–99. doi: 10.1006/jesp.2001.1490

Gilovich, T., Medvec, V. H., and Savitsky, K. (2000). The spotlight effect in social judgment: an egocentric bias in estimates of the salience of one's own actions. *J. Pers. Soc. Psychol.* 78, 211–222. doi: 10.1037/0022-3514.78.2.211

Gotlib, I. H., and Joormann, J. (2010). Cognition and depression: current status and future directions. *Annu. Rev. Clin. Psychol.* 6, 285–312. doi: 10.1146/annurev.clinpsy.121208.131305

Guilbault, R. L., Bryant, F. B., Brockway, J. H., and Posavac, E. J. (2004). A meta-analysis of research on hindsight bias. *Basic Appl. Soc. Psychol.* 26, 103–117. doi: 10.1207/s15324834basp2602andamp;3_1

Haddock, T. B., and Birch, S. A. J. (2024). The relationship between children's theory of mind and social-emotional health. *Psychol. J. Res. Open* 6, 1–15. doi: 10.31038/PSYJ.2024613

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *J. Econ. Behav. Organ.* 53, 3–35. doi: 10.1016/S0167-2681(03)00094-5

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., and Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317, 1360–1366. doi: 10.1126/science.1146282

Hiemstra, W., Orobio de Castro, B., and Thomaes, S. (2018). Reducing aggressive children's hostile attributions: a cognitive bias modification procedure. *Cogn. Ther. Res.* 43, 387–398. doi: 10.1007/s10608-018-9958-x

Hinds, P. J. (1999). The curse of expertise: the effects of expertise and debiasing methods on predictions of novice performance. *J. Exp. Psychol. Appl.* 5, 205–221. doi: 10.1037//1076-898X.5.2.205

Hirsch, C. R., Krahé, C., Whyte, J., Loizou, S., Bridge, L., Norton, S., et al. (2018). Interpretation training to target repetitive negative thinking in generalized anxiety disorder and depression. *J. Consult. Clin. Psychol.* 86, 1017–1030. doi: 10.1037/ccp0000310

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., and Fang, A. (2012). The efficacy of cognitive behavioral therapy: a review of meta-analyses. *Cogn. Ther. Res.* 36, 427–440. doi: 10.1007/s10608-012-9476-1

Hooper, N., Erdogan, A., Keen, G., Lawton, K., and McHugh, L. (2015). Perspective taking reduces the fundamental attribution error. *J. Contextual Behav. Sci.* 4, 69–72. doi: 10.1016/j.jcbs.2015.02.002

Hotte-Meunier, A., Penney, D., Mendelson, D., Thibaudeau, É., Moritz, S., Lepage, M., et al. (2024). Effects of metacognitive training (MCT) on social cognition for schizophrenia spectrum and related psychotic disorders: a systematic review and meta-analysis. *Psychol. Med.* 54, 914–920. doi: 10.1017/S0033291723002611

Hubbard, E. G. (1911). *A Thousand and One Epigrams, Selected from the Writings of Elbert Hubbard*. Published by East Aurora, New York: The Roycrofters.

Hughes, C., and Ensor, R. (2006). Behavioural problems in 2-year-olds: links with individual differences in theory of mind, executive function and harsh parenting. *J. Child Psychol. Psychiatry* 47, 488–497. doi: 10.1111/j.1469-7610.2005.01519.x

Hughes, C., and Ensor, R. (2007). Positive and protective: effects of early theory of mind on problem behaviors in at-risk preschoolers. *J. Child Psychol. Psychiatry* 48, 1025–1032. doi: 10.1111/j.1469-7610.2007.01806.x

Imuta, K., Henry, J. D., Slaughter, V., Selcuk, B., and Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: a meta-analytic review. *Dev. Psychol.* 52, 1192–1205. doi: 10.1037/dev0000140

Kahneman, D, and ProQuest (Firm). (2011). *Thinking, Fast and Slow*. Toronto, ON: Farrar, Straus and Giroux.

Keysar, B., Lin, S., and Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition* 89, 25–41. doi: 10.1016/S0010-0277(03)00064-7

Kindt, M., and Van Den Hout, M. (2001). Selective attention and anxiety: a perspective on developmental issues and the causal status. *J. Psychopathol. Behav. Assess.* 23:193. doi: 10.1023/A:1010921405496

Livingston, L. A., Shah, P., and Happé, F. (2024). Linearly integrating speed and accuracy to measure individual differences in theory of mind: evidence from autistic and neurotypical adults. *Q. J. Exp. Psychol.* 77, 287–297. doi: 10.1177/17470218231165251

Macrae, C. N., Mitchell, J. P., McNamara, D. L., Golubickis, M., Andreou, K., Møller, S., et al. (2016). Noticing future me. *Personal. Soc. Psychol. Bull.* 42, 855–863. doi: 10.1177/0146167216644961

Mathews, A., and Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *J. Abnorm. Psychol.* 109, 602–615. doi: 10.1037/0021-843X.109.4.602

Mathews, A., and MacLeod, C. (2002). Induced processing biases have causal effects on anxiety. *Cogn. Emot.* 16, 331–354. doi: 10.1080/02699930143000518

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., and Kassam, K. S. (2015). Debiasing decisions: improved decision making with a single training intervention. *Policy Insights Behav. Brain Sci.* 2, 129–140. doi: 10.1177/2372732215600886

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: imputing one's own knowledge to others. *Psychol. Bull.* 125, 737–759. doi: 10.1037/0033-2909.125.6.737

Nisbett, R. E., and Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Oeberst, A., and Imhoff, R. (2023). Toward parsimony in bias research: a proposed common framework of belief-consistent information processing for a set of biases. *Perspect. Psychol. Sci.* 18, 1464–1487. doi: 10.1177/17456916221148147

Peterson, C., Slaughter, V., Moore, C., and Wellman, H. M. (2016). Peer social skills and theory of mind in children with autism, deafness, or typical development. *Dev. Psychol.* 52, 46–57. doi: 10.1037/a0039833

Peterson, J. L., Bellows, A., and Peterson, S. (2015). Promoting connection: perspective-taking improves relationship closeness and perceived regard in participants with low implicit self-esteem. *J. Exp. Soc. Psychol.* 56, 160–164. doi: 10.1016/j.jesp.2014.09.013

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/S0140525X00076512

Rector, N. A., and Beck, A. T. (2002). Cognitive therapy for schizophrenia: from conceptualization to intervention. *Can. J. Psychiatry.* 47, 39–48. doi: 10.1177/070674370204700107

Repacholi, B. M., and Slaughter, V. (2003). *Individual Differences in Theory of Mind: Implications for Typical and Atypical Development*. New York, NY: Psychology Press.

Ross, L., Greene, D., and House, P. (1977). The false consensus effect: an egocentric bias in social perception and attribution processes. *J. Exp. Soc. Psychol.* 13, 279–301. doi: 10.1016/0022-1031(77)90049-X

Samson, C., Livet, A., Gilker, A., Potvin, S., Sicard, V., and Lecomte, T. (2024). Reasoning and interpretation cognitive biases related to psychotic characteristics: an umbrella-review. *PloS ONE* 19:e0314965. doi: 10.1371/journal.pone.0314965

Savitsky, K., Keysar, B., Epley, N., Carter, T., and Swanson, A. (2011). The closeness-communication bias: increased egocentrism among friends vs. strangers. *J. Exp. Soc. Psychol.* 47, 269–273. doi: 10.1016/j.jesp.2010.09.005

Shih, M., Wang, E., Trahan Bucher, A., and Stotzer, R. (2009). Perspective taking: reducing prejudice towards general outgroups and specific individuals. *Group Process. Intergroup Relat.* 12, 565–577. doi: 10.1177/1368430209337463

Slaughter, V., Imuta, K., Peterson, C. C., and Henry, J. D. (2015). Meta-analysis of theory of mind and peer popularity in the preschool and early school years. *Child Dev.* 86, 1159–1174. doi: 10.1111/cdev.12372

Smith, K. P. (2017). Bullying and theory of mind: a review. *Curr. Psychiatry Rev.* 13, 90–95. doi: 10.2174/1573400513666170502123214

Sutherland, S. L., and Cimpian, A. (2015). Children show heightened knew-it-all-along errors when learning new facts about kinds: evidence for the power of kind representations in children's thinking. *Dev. Psychol.* 51, 1115–1130. doi: 10.1037/a0039463

Taylor, M., Esbensen, B. M., and Bennett, R. T. (1994). Children's understanding of knowledge acquisition: the tendency for children to report that they have always known what they have just learned. *Child Dev.* 65, 1581–1604. doi: 10.2307/1131282

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131. doi: 10.1126/science.185.4157.1124

van Brussel, S., Timmermans, M., Verkoeijen, P., and Paas, F. (2021). Teaching on video as an instructional strategy to reduce confirmation bias—a pre-registered study. *Instr. Sci.* 49, 475–496. doi: 10.1007/s11251-021-09547-4

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128. doi: 10.1016/0010-0277(83)90004-5

Wolgast, A., Tandler, N., Harrison, L., and Umlauft, S. (2020). Adults' dispositional and situational perspective-taking: a systematic review. *Educ. Psychol. Rev.* 32, 353–389. doi: 10.1007/s10648-019-09507-y

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 04 July 2025
DOI 10.3389/fpsyg.2025.1546464

Check for updates

OPEN ACCESS

EDITED BY
Sara Isernia,
Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy

REVIEWED BY
Schnell Zsuzsanna,
University of Pécs, Hungary
Giulia Funghi,
University of Trento, Italy

*CORRESPONDENCE
Carla Sharp
csharp2@uh.edu

RECEIVED 16 December 2024
ACCEPTED 12 June 2025
PUBLISHED 04 July 2025

CITATION
Sharp C, Barr C and Vanwoerden S (2025) Hypermentalizing: the development and validation of a self-report measure. *Front. Psychol.* 16:1546464.
doi: 10.3389/fpsyg.2025.1546464

Hypermentalizing: the development and validation of a self-report measure

Carla Sharp[1]*, C. Barr[2] and Salome Vanwoerden[3]

[1]Department of Psychology, University of Houston, Houston, TX, United States, [2]STEM Programs, Rice University, Houston, TX, United States, [3]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States

Introduction: Hypermentalizing (referred to as excessive theory of mind or biased mindreading) is defined as the tendency to make assumptions about other people's mental states that go beyond observable data. Despite recent interest in this construct, no self-report measure of hypermentalizing exists. The aim of the current study was to fully operationalize the construct of hypermentalizing by developing a theoretically grounded (attachment-based) self-report measure of hypermentalizing assessing mentalizing related to parents, peers and intimate partners; and evaluate the new measure for its psychometric properties.

Methods: In Study 1,745 undergraduate students (mean age 21.12; *SD* = 2.19) completed the Hypermentalizing Questionnaire (HMZQ) alongside an experimental measure of mentalizing (the Movie Assessment for Social Cognition; MASC).

Results: Results of factor analyses with MASC scores for external validity confirmed the purported factor structure of the HMZQ and suggested superiority for the HMZQ version that assesses mentalizing in relation to parents. Study 2 compared HMZQ scores in 364 adolescents between 12 and 17 years of age (70 adolescents with BPD, 136 psychiatric controls, and 158 healthy controls), and confirmed the superiority of the 26-item version of the HMZQ that assesses mentalizing in relation to parents, in that it was only the HMZQ version that distinguished borderline personality disorder from other psychiatric disorders and healthy controls.

Discussion: The current study provides evidence in support of the HMZQ to assess hypermentalizing in typical and atypical populations of adolescents and young adults.

KEYWORDS

hypermentaling, personality disorder, assessment, psychopathology, theory of mind

Introduction

Mentalizing is a multi-component construct defined as the capacity to reflect on one's own thoughts and feelings and those of others to predict and understand behavior in the context of interpersonal interactions and relationships (Bateman and Fonagy, 2012a,b). The concept has been used in psychoanalytic literature since the 1970s (Allen, 2003; Marty, 1991; Marty and M'Uzan, 1963) to refer to the process of mental elaboration, including symbolization for the transformation and elaboration of drive-affect experiences as mental phenomena and structures (Lecours and Bouchard, 1997). It was incorporated into mainstream neurobiological and developmental literature (Frith, 1992; Morton, 1989) in the 1980s and 1990s, where it has been used interchangeably with the more frequently used concept of "theory of mind" (ToM). Premack and Woodruff (1978) coined the term "theory of mind" to refer to the capacity to interpret other people's behavior within a mentalistic framework in order to understand how self and others think, feel, perceive, imagine, react, attribute, infer, and so on. Mentalizing lies at the very core of our humanity because without the capacity to reflect on our own and other's mental states, we cannot maintain constructive social interaction, mutuality in relationships, or a robust and integrated sense of self (Bateman and Fonagy, 2012a,b).

The capacity to mentalize is theorized to develop within the context of secure attachment relationships with primary caregivers (Fonagy et al., 2002). Empirical studies support this notion with prospective studies having demonstrated that secure attachment facilitates the development of mentalizing (Fonagy et al., 1991; Meins, 1998; Symons and Clark, 2000). Conversely, disruptions in attachment relationships are associated with impaired mentalizing capacity, both prospectively (Belsky et al., 2012) and cross-sectionally (Sharp et al., 2015b,a; Ensink et al., 2015). In turn, impairment in mentalizing has been demonstrated for almost all types of psychopathology in youth (see Sharp and Venta, 2012 for a review) and adults (see Brune and Brune-Cohrs, 2006 for a review). Emerging from this literature are two broad types of mentalizing impairment: hypomentalizing and hypermentalizing (Abu-Akel, 2008; Crespi and Badcock, 2008; Fonagy et al., 2016; Gambin et al., 2015). *Hypomentalizing* reflects a deficiency in (lack of) mentalizing; that is, an inability to consider complex models of one's own mind and/or that of others (Fonagy et al., 2016). This deficiency is likely due to a reduced capacity to attribute thoughts, feelings and intentions (i.e., mental states) to oneself and others, resulting in comprised ability to make sense of social cues and interpersonal interactions. A large body of literature demonstrates an association between hypomentalizing and a wide variety of disorders, including autism (e.g., Baron-Cohen, 2000), psychopathy (e.g., Sharp et al., 2015a), and conduct problems (e.g., Happé and Frith, 1996; Sharp, 2008).

In contrast, *hypermentalizing*, which has also been referred to as excessive theory of mind (Dziobek et al., 2006) or biased mindreading (Sharp, 2000), involves making assumptions about other people's mental states that go beyond observable data (Crespi and Badcock, 2008; Fonagy et al., 2016; Gambin et al., 2015; Sharp, 2014; Sharp et al., 2011; Sharp and Vanwoerden, 2015). As such, it involves over-attribution of mental states and intentions to others, their likely misinterpretation, and the urge to act in response to the assumed mental states of others. It furthermore involves the over-interpretation of one's own mental states and a conflation of self-other mental states (Frick et al., 2012) or overactive and exaggerated resonance with the mental states of others due to confusion between self-and other-mental states (Ensink et al., 2015; Sharp and Vanwoerden, 2015). Hypermentalizing is by its very nature indicative of a metacognitive deficit since it involves failure to attain a higher-order representation from which to question one's own belief in service of generating alternative hypotheses in interpreting situations about the self and others (Semerari et al., 2005, 2007). As such, hypermentalizing reflects a lack of metacognitive differentiation (Semerari et al., 2005), because representation is conflated with reality. In summary then, hypomentalizing represents deficient (under) use of mental states in explaining behavior in self and others, while hypermentalizing represents over-use of mental states in making sense of self and others. In contrast, optimal mentalizing entails the use of mental states to understand self and others in productive ways. For instance, the optimal mentalizer would use mental states to explain behavior, but would do so from a stance of curiosity, openness and flexibility. The optimal mentalizer would ask whether a feeling, thought or intention is associated with behavior, but would not assume such mental states. Finally, the optimal mentalizer is able to flexibly integrate new information as a representation of another's (or own) mind is constructed in service of understanding and explaining behavior (Sharp and Bevington, 2022).

Compared to the evidence base on hypomentalizing in psychopathology, the hypermentalizing literature is much smaller, which is partly due to a lack of measures to assess the construct. Hypermentalizing first appeared in the literature in the context of schizophrenia (Langdon and Coltheart, 1999; Langdon et al., 2006a,b). Patients with schizophrenia have been found to overattribute intentions; misplace emphasis on stimuli thereby prompting inferences of abnormal meaning, see patterns that other people do not perceive, draw conclusions on less information, and report false-positives in ambiguous situations (Abu-Akel and Shamay-Tsoory, 2011; Grant et al., 2014; Howes and Kapur, 2009). More recently, hypermentalizing is also assessed as a key feature in Borderline Personality Disorder (BPD; Bo et al., 2015; Franzen et al., 2011; Frick et al., 2012; Preissler et al., 2012; Sharp et al., 2011), and can be detected most clearly in long and overly detailed accounts that have little or no relationship to reality, coupled with inflexible certainty in beliefs about others' mental states (Fonagy et al., 2016).

The increase in interest in the construct of hypermentalizing calls for the development of reliable and valid tools for its assessment. The most commonly used tool is the Movie Assessment for Social Cognition (MASC; Dziobek et al., 2006) which is an experimental task in which research participants are presented with four mutually exclusive options in response to a video clip of interaction partners: hypermentalizing, hypomentalizing, no mentalizing and accurate mentalizing. While this task has been shown to be valid and reliable in many studies across various populations, it is time consuming and can take up to 45 min to complete. As yet, a relatively quick and easy-to-administer self-report tool for the assessment of hypermentalizing is lacking. Apart from the practical advantages associated with the brevity and ease of administration of self-report, an additional advantage relates to the fact that the MASC is a performance-based measure. It is well-known that performance-based measures tap into one aspect of a construct while self-report measures tap into the more conscious, representational aspects of the construct. Both are important and provide important insight into mentalizing through different lenses.

The aim of the current study was to develop a self-report measure of hypermentalizing and evaluate the newly developed measure for its psychometric properties. In the development and evaluation of the measure, a few considerations were taken into account. First, a review of the hypermentalizing literature revealed that hypermentalizing contains five elements (Sharp, 2014; Sharp and Vanwoerden, 2015): an overconcern with the mental states of others; overinterpretation of others' mental states; inflexible certainty in own beliefs about others' mental states; acting impulsively on the assumed mental states of others; and second-guessing or over-interpretation of own mental states. These related components were identified as forming potential subscales. Items were subsequently written for each.

Second, given that the concept of mentalizing, at least in the context of personality disorder, has its roots in attachment, our intention was to develop a measure of hypermentalizing that acknowledges this theoretical basis. Specifically, Fonagy and colleagues' model for the development of personality disorder suggest that it is through secure attachment with caregivers and its associated parental reflective functioning that the mentalizing capacity of the child emerges (Fonagy et al., 2002; Sharp and Fonagy, 2008). In short, infants and young children do not yet have the reflective capacity to help them make sense of self and others. The development of mentalizing capacity therefore relies on a process called "marked mirroring" by which caregivers mark their offspring's internal experiences and give it back to the offspring in digested form. In this way, over time, and as the child's own reflective capacities increase, the infant/child comes to know their own mental states and develop a capacity for reading the mental states of others. If caregivers are intrusive in their marked mirroring or passive, or inconsistent, or non-contingent, atypical mentalizing styles, including hypo- and hypermentalizing develop (see e.g., Kim, 2015 for a review). Against the background of this theoretical and empirical evidence, we deemed it important to relate items of the newly developed mentalizing measure directly to individuals' attachment context. Participants were therefore asked to answer questions about thoughts and feelings that are typical for them in interaction with significant others. Relatedly, against the background of research showing attachment and mentalizing is relationship-specific (O'Connor and Hirsch, 1999), it was important that the measure be sensitive to the specific attachment context. Therefore, questions were asked three times over: in relation to parents, romantic partners and closest friend. These three attachment contexts were chosen because of their unique relevance for the developmental period of adolescence and young adulthood. Specifically, research has shown that adolescents typically increase their valuation of peer and romantic partner relationships, develop greater psychological distance from parents, and renegotiate boundaries and responsibilities in family relationships (Fuligni and Eccles, 1993; Hallquist et al., 2015; Steinberg et al., 2006). While research has shown that the quality of the parent-child attachment relationship tends to influence the quality of peer and romantic partner attachment via internal working models that establish patterns of interpersonal relationships, it is also true that significant changes are made in the organization of attachment systems during adolescence and young adulthood such that the correlation between parent, peer and romantic partner attachment may diminish in some individuals (Gorrese and Ruggieri, 2012), resulting in attachment and mentalizing that are relationship-specific (O'Connor and Hirsch, 1999). Having questions asked for three attachment contexts meant that data analytic strategies had to take into account method factors influencing response patterns. It also raised the interesting question whether context-specific hypermentalizing probing was even necessary. Put differently, do factors that are comprised of shared variability of these three perspectives (parent, romantic partner, and closest friend) provide any incremental information beyond what would be obtained from the total score when responses across the three attachment context were summed? Additionally, it would be important to determine which of the three relationships contexts provided the most useful information about hypermentalizing compared to the others.

Third, given that hypermentalizing has been demonstrated in adolescent and adult populations using the MASC, and to facilitate studies in which the development of hypermentalizing can be tracked, it was desirable to develop items that could be used in adolescent and young adult samples.

To this end, we conducted two studies, each with a unique age cohort that covers the developmental period during which attachment begins to transition to include peer and romantic partners—that is, adolescence and young adulthood. Consistent with the World Health Organization's definition of the term "young people" to denote 10–24 year-olds, this developmental period extends from puberty (operationally defined as age 10–12 years), beyond traditional notions of adolescence (ending at age 18 years), to around 25 years of age (Dahl et al., 2018; Sawyer et al., 2018). Study 1 made use of a convenience sample of young adults in a college setting and its primary focus was twofold; first, to evaluate the factor structure of the newly developed measure, and second, to evaluate the association with the MASC as a criterion measure of hypermentalizing. Study 2 utilized an adolescent sample consisting of three groups: typically developing adolescents, adolescents with borderline pathology and adolescents with psychiatric problems (but no borderline pathology). The aim of Study 2 was to conduct a three-group comparison to give credence to the psychopathology roots of the hypermentalizing construct, and its particular significance for borderline personality pathology. For instance, individuals with social anxiety have also been shown to hypermentalize (Hezel and McNally, 2014; Washburn et al., 2016). A recent meta-analysis furthermore showed that hypermentalizing was not specific to BPD (McLaren et al., 2022). We hypothesized, however, that hypermentalizing would be more profoundly affected in borderline personality disorder compared to psychiatric caseness in general.

Study 1: factor structure and associations with a criterion measure

Study 1 utilized a large college-based sample which afforded us the opportunity to explore two aims. First, we evaluated the factor structure of the measure given the increased variability of most constructs in non-clinical samples. In this study we were interested in determining the best model to explain covariance between items to ultimately justify the use of the parent-, romantic partner- and close friends versions of the measure, as well as the five subscales representing five underlying factors (an overconcern with the mental states of others; overinterpretation of others' mental states; inflexible certainty in own beliefs about others' mental states; acting impulsively on the assumed mental states of others; and second-guessing or over-interpretation of own mental states) of hypermentalizing. Second, we evaluated the associations with the gold standard measure of hypermentalizing, namely the MASC (Dziobek et al., 2006).

Study 1 methods

Participants

Data were collected from a sample of 745 undergraduate students at a large and racially and ethnically diverse university in an urban area in the southern region of the United States. Participants were recruited via a mass email advertising an online study to undergraduate students enrolled in at least one Psychology course. The recruitment email was sent from the Department of Psychology and participants self-selected to participate in this study by following a hyperlink to the University's online survey system. All responses were anonymous and identifiable only by a unique, randomly generated code. Inclusion criteria were English fluency and age between 18 and 25. There were no exclusion criteria. Participants were informed of the inclusion criteria in a cover letter and were instructed to self-exclude if the aforementioned criteria were not met. The sample included 586 women and 159 men (9 participants did not identify their gender). The mean age in this sample was 21.12 ($SD = 2.19$). The self-identified ethnoracial breakdown was 26.1% White/Not Hispanic, 14.7% Black, 31.5% Hispanic/Not White, 23.2% Asian or Pacific Islander, and 4% Multiracial or other (eight participants did not identify their ethnoracial background). This study was approved by the relevant Institutional Review Board and informed consent was provided. Participants completed questionnaires via a web-based program and were compensated with research credit.

Measures

Hypermentalizing questionnaire (HMZQ)

The HMZQ consists of 26 items that are completed on a five-point Likert scale ranging from 0 ("not typical at all") to 4 ("very typical). Respondents are asked to do the following: "*Below are 26 questions about the thoughts and feelings that are TYPICAL for you in interaction with your SIGNIFICANT OTHERS. We will ask the questions three times over. First, in the context of your relationships with your parents. Second, in the context of the relationship with your romantic partner (if you are currently not in a relationship, think back to your most recent relationship). And third, in the context of the relationship with your closest friend.*" Items were written to load onto one of five underlying factors: Overconcern for the mental state of others (OC), e.g., "I worry a lot about what my parents are thinking and feeling"; Overinterpretation of others' mental states (OI), e.g., "My parents often say I overinterpret interpersonal situations with them"; Inflexible certainty in beliefs about others' mental states (IC), e.g., "My feelings about what my parents are thinking are hardly ever wrong"; Acting impulsively on assumed mental states of others (IP), e.g., "I easily lose control in situations with my parents if my feelings get hurt"; Second-guessing/over-interpretation of own mental states (SG), e.g., "I often second-guess myself when interacting with my parents." These components of hypermentalizing were derived from a thorough literate review of hypermentalizing (Sharp, 2014; Sharp and Vanwoerden, 2015). Items were written during a workshop of experts and piloted for among undergraduates in the first author's lab. Items were refined based on informal feedback.

Movie assessment for social cognition (MASC)

To explore criterion validity, participants completed the MASC (Dziobek et al., 2006), which is a computerized test for the assessment of mentalizing that approximates the demands of everyday life. Subjects watch a 15-min film about four characters getting together for a dinner party. Themes of each segment covered friendship and dating issues. During administration of the task, the film is stopped at 45 points and multiple-choice questions referring to the characters' mental states (feelings, thoughts, and intentions) are asked (e.g., "What is Betty feeling?", "What is Cliff thinking?"). All items answered correctly are summed for a total score with higher scores indicating higher mentalizing capacity. The three incorrect responses are categorized as representing hypermentalizing, undermentalizing, or no mentalizing; counts of each of these incorrect responses make up the subscales of maladaptive mentalizing. The MASC is a reliable instrument that has proven sensitive in detecting subtle mindreading difficulties in adults of normal IQ (Dziobek et al., 2006) and in adolescents (Sharp et al., 2011).

Data analytic strategy

To investigate the aims of study 1, we examined (1) competing factor models to determine which model or models were a reasonable fit to the data using four competing factor models described below, (2) if HMZQ trait factors accounted for unique criterion validity beyond total scores of the three types of relationships (parent close friend, and romantic partner) in higherarchical regression models, and (3) convergent and discriminant validity patterns between HMZQ trait factors, HMZQ total scores, and MASC dimensions, by examining the zero-order correlations between HMZQ factors/total scores and MASC dimensions.

To investigate the factor structure of the newly developed HMZQ, we investigated four competing models. All four models were non-nested so we relied on traditional fit indices: Comparative Fit Index (CFI), Tucker-Lewis Index (TLI) and Root Mean Square Error of Approximation (RMSEA). Item variance accounted to evaluate the best fitting model. The four competing models were (1) a single trait model with all items loading onto a hypermentalizing factor, (2) a single trait multimethod model with all items loading onto a hypermentalizing factor and the respective items loading onto parent, romantic partner, and close friend method factors, (3) a multi trait model with items loading onto OC, OI, IF, IP, and SG factor, and (4) a multi trait multimethod model with all items loading OC, OI, IF, IP, and SG trait factors and the respective items loading onto parent, romantic partner, and close friend method factors.

Study 1 results

Factor structure of the HMZQ

Model fit statistics are presented in Table 1.

As can be seen, models 1 and 3, which did not include method factors did not fit well, but both models 2 and 4 did fit well. Model 4, the multi-trait multi-method model fit the best, but it's

TABLE 1 Hypermentalizing factor model fit indices.

Models tested	CFI	TLI	RMSEA
Model 1: single trait Unidimensional hypermentalizing trait with no modeling of the attachment context	0.761	0.754	0.085
Model 2: single trait, multi-method Unidimensional hypermentalizing trait with modeling of the attachment context	0.946	0.943	0.041
Model 3: multi-trait Hypermentalizing components modeled with no modeling of the attachment context	0.764	0.757	0.085
Model 4: multi-trait, multi-method Hypermentalizing components model with modeling of attachment context	0.948	0.945	0.040

To interpret model fit statistics for RMSEA, values of 0.05 and lower are generally good, while those between 0.05 and 0.08 are acceptable. For CFI and TLI values close to or higher than 0.09 are general good.

improvement over model 2, the single trait multi-method model was only marginal. Thus, for reasons of parsimony, it could be argued that model 2 is the preferable model. However, we retained both models 2 and 4 in our investigation of the best predictor of performance on the MASC. Finally, we also investigated factor models containing only parent items, only close friend items, and only romantic partner items. However, these factor scores correlated at or above 0.94 with their respective total scores. Thus, because the factor scores are near replications of the total scores, we dropped these factors from our predictive models of the MASC as a practical consideration of field use of the instrument.

Relations between HMZQ and MASC

Descriptive statistics

Descriptive statistics and intercorrelations between the four subscales of the MASC (hypermentalizing, hypomentalizing, no mentalizing, and correct mentalizing), the HMZQ trait factor, the five HMZQ trait factors from the five factor model, the HMZQ total score, the HMZQ parent total, the HMZQ romantic partner total, and the HMZQ close friend total can be seen in Table 2.

Predictors of MASC hypermentalizing

To evaluate criterion validity, we sought to investigate both which variables predicted hypermentalizing on the MASC and if the factors identified in our two factor models accounted for additional variance beyond the HMZQ total scores. As seen in Table 2, all variables except the romantic partner total score have significant zero order correlations with MASC hypermentalizing. From a purely descriptive standpoint, the magnitude of the single trait factor had the highest correlation with MASC hypermentalizing.

To investigate if the trait factors from the single and multi-trait models accounted for incremental variability in MASC

TABLE 2 Descriptive statistics.

Obs	Variable	Mean	Std. Dev.	1	2	3	4	6	7	8	9	10	11	12	13
1	MASCTotal	17.19	15.65	–	–	–	–	–	–	–	–	–	–	–	–
2	MASCExcessiveToM	4.00	4.24	0.48**	–	–	–	–	–	–	–	–	–	–	–
3	MASCLessToM	3.23	3.83	0.36**	0.69**	–	–	–	–	–	–	–	–	–	–
4	MASCNoToM	1.93	2.99	0.17**	0.61**	0.76**	–	–	–	–	–	–	–	–	–
6	Overconcern others (OC)	0.01	0.84	0.37**	−0.25**	−0.39**	−0.35**	–	–	–	–	–	–	–	–
7	Over interpretation others (OI)	0.01	0.88	0.36**	−0.26**	−0.36**	−0.34**	0.7**	–	–	–	–	–	–	–
8	Inflexible certainty others (IC)	0.00	0.88	0.37**	−0.26**	−0.35**	−0.33**	0.8**	0.76**	–	–	–	–	–	–
9	Acting impulsively (IP)	0.01	0.91	0.36**	−0.26**	−0.34**	−0.32**	0.64**	0.98**	0.77**	–	–	–	–	–
10	Second-guessing self (SG)	0.01	0.95	0.38**	−0.28**	−0.36**	−0.34**	0.74**	0.91**	0.92**	0.94**	–	–	–	–
11	Grand_Total_Hypermentalizing	122.15	49.03	−0.11**	0.14**	0.05	0.07	0.05	0.03	0.04	0.02	0.03	–	–	–
12	Total_HMZ_Parents	41.49	18.98	−0.11**	0.14**	0.02	0.07	0.09*	0.08*	0.1**	0.07	0.08*	0.88**	–	–
13	Total_HMZ_Romantic	46.10	18.44	0.02	0.03	−0.05	−0.03	0.21**	0.23**	0.2**	0.21**	0.21**	0.85**	0.63**	–
14	Total_HMZ_ClosestFriend	34.95	19.58	−0.25**	0.22**	0.18**	0.2**	−0.14**	−0.18**	−0.14**	−0.17**	−0.16**	0.87**	0.66**	0.59**

* < .05, ** < .01.

TABLE 3 Hierarchical regressions.

Variable	b1	SE1	b2	SE2	b3	SE3	R^2
Model 1							
Parent HMZ total	0.02*	0.01	0.01	0.01	0.01	0.01	
Set 1							0.015
Romantic partner HMZ total			−0.03*	0.01	−0.02	0.01	
Close friend HMZ total			0.05*	0.01	0.03*	0.01	
Set 2							0.060
Hypermentalizing trait factor					0.91*	0.16	
Set 3							0.108
Model 2							
Parent HMZ total	0.02*	0.01	0.01	0.01	0.02	0.01	
Set 1							0.015
Romantic partner HMZ total			−0.03*	0.01	−0.02	0.01	
Close friend HMZ total			0.05*	0.01	0.03*	0.01	
Set 2							0.060
OC trait factor					−0.51	0.37	
OI trait factor					2.07	1.38	
IF trait factor					−0.96	0.76	
IP trait factor					−3.30	2.03	
SG trait factor					1.50	1.51	
Set 3							0.111

hypermentalizing beyond the parent, romantic partner, and close friend total scores, we examined two hierarchical regression models (1) a model with three sets of variables with set 1 consisting of the parent total score component of the HMZQ, set 2 consisting of the romantic partner and close friend total scores from the HMZQ, and set 3 being comprised of the hypermentalizing trait factor from the single factor multi method model, and (2) a model with the same two first sets as the previous model but where set 3 consists of the five trait factors from the multi trait multi method model. We chose this set because of past research suggesting that the parental relationship was the most well established relationship of the three types of attachment context given its longer duration. Additionally, the factor scores were entered last to determine if complex modeling, something many practitioners may not undertake when collecting client data, improved variance accounted in MASC hypermentalizing. The results of these two hierarchical regressions can be seen in Table 3.

In both hierarchical models all predictor sets accounted for significant incremental variance accounted in MASC hypermentalizing. These results suggest that the hypermentalizing factors identified in the factor modeling do account for incremental validity in MASC mentalizing. It should be noted that both hierarchical models accounted for comparable variability in MASC hypermentalizing (a difference of only 0.003) and thus, these results provide further support for the utility of the single factor hypermentalizing trait factor.

Correlations with MASC subscales

Finally, we investigated the patterns of association for the HMZQ total scores for all four components of the MASC: hypermentalizing, hypomentalizing, no mentalizing and accurate mentalizing. To study patterns, we first examined the patterns of significance in the zero order correlations. We noted that the HMZQ trait factor not only correlated with MASC hypermentalizing, but also with hypomentalizing and no mentalizing. The same was true for the five trait factors, except in the opposite direction. As mentioned in the description of the MASC, scores are somewhat dependent on each other—the higher the scores on any incorrect response (hyper-, hypo-, no-) the lower the total correct on the MASC. Thus, given that the pattern of correlations between the both the single trait factor and the five trait factors with hypermentalizing, hypomentalizing, no mentalizing were consistent, the results suggest that the HMZQ trait factor is measuring incorrect mentalizing regardless of the error (excessive/no/less). Likewise, all five trait factors seem to be measuring accurate mentalizing. The close friend total score seemed to follow a similar pattern to the single HMZQ trait factor. The romantic partner total score was unrelated to any facet of the MASC. Finally, the parent total related to MASC hypermentalizing but not to hypo- or no mentalizing. Thus, from these results, only the parent total score seems to discriminate between MASC hypermentalizing and hypo- and no mentalizing.

Discussion study 1

The aim of Study 1 was to evaluate whether the purported factor structure of the HMZQ is supported. To this end, we examined four factor models and factor scores were evaluated for their associations with a criterion measure of mentalizing (the MASC). Results of the factor analyses demonstrated support for both a single trait multimethod model with all items loading onto a hypermentalizing factor and the respective items loading onto parent, romantic partner, and close friend method factors, as well as a multi trait multimethod model with all items loading onto the OC, OI, IF, IP, and SG trait factors and the respective items loading onto parent, romantic partner, and close friend method factors. To further explore the utility of each of these models, they were evaluated in regression analyses with MASC hypermentalizing as dependent variable and the factor scores of both models as predictor variables. Results of the regression analyses showed that models with both the single trait factor, as well as the multiple traits accounted for significant incremental validity in MASC hypermentalizing beyond the parent, close friend, and romantic partner total scores. Additionally, the two models were comparable in explaining variability in MASC hypermentalizing suggesting the utility of the single factor hypermentalizing trait factor as most parsimonious. Evaluation of the associations between method and HMZQ factors and MASC subscales demonstrated that the HMZQ factors were correlated with all MASC subscales, not just the hypermentalizing subscale of the MCAS. In this sense, the HMZQ factors provide convergent validity for MASC hypermentalizing, but not discriminant validity with hypomentalizing, no mentalizing, and accurate mentalizing. As such, the HMZQ factors may be more reflective of accurate mentalizing and not specifically to the error of hypermentalizing. On the other hand, the parent total score did follow the expected pattern with regard to convergent and discriminant validity with the parent total score relating to negatively with the MASC total, positively with hypermentalizing, and not significantly relating to hypo or no mentalizing. In this sense, the parent total score of the measure appears to be best at discriminating between different forms of mentalizing, although we do note that the trait factors do account for significant incremental validity in terms of prediction of MASC hypermentalizing.

Study 2: clinical utility of the HMZQ

The aim of Study 2 was to investigate clinical utility. To this end, Study 2 made use of a clinical sample of adolescents who were well-characterized psychiatrically to derive a borderline vs. non-borderline psychiatric group, as well as a sample of typically developing adolescents recruited from the community. We specifically chose BPD as a comparison because the concept of mentalizing as used in psychotherapy originated in the context of BPD (Fonagy, 1991) and hypermentalizing was identified as a potentially unique correlate of BPD in its early conceptualization (Sharp et al., 2011; Sharp, 2014). To assess BPD, we use an interview-based measure of BPD so as to reduce shared method variance between the psychopathology measure (BPD) and the new developed self-report measure of hypermentalizing.

Study 2 methods

Participants

Participants for Study 2 included 320 adolescents who were recruited from a psychiatric inpatient unit that serves individuals with severe behavioral and emotional disorders. Of these adolescents, 97 met full criteria for borderline personality disorder (BPD) as determined by clinical interview (the CI-BPD; Zanarini et al., 2003). Additionally, 189 healthy controls were recruited from the community through schools and community programs. Inclusion criteria for both samples was sufficient proficiency in English to consent and complete the necessary assessments, and exclusion criteria were a diagnosis of schizophrenia or another psychotic disorder, an autism spectrum diagnosis, or an IQ of <70. A number of adolescents did not complete the hypermentalizing questionnaire (29 healthy controls, 68 psychiatric controls, and 27 with BPD); however, these adolescents did not differ from those who did complete the questionnaire in terms of age [healthy controls: $t_{(185)} = -0.84$, $p = 0.40$; psychiatric controls: $t_{(202)} = -0.14$, $p = 0.89$; BPD $t_{(95)} = -0.50$, $p = 0.62$] or gender [healthy controls: $\chi^2_{(1,188)} = 2.21$, $p = 0.14$; psychiatric controls: $\chi^2_{(1,204)} = 0.04$, $p = 0.84$; BPD $\chi^2_{(1,97)} = 0.33$, $p = 0.56$]. Therefore, the final sample included 70 adolescents with BPD, 136 psychiatric controls, and 158 healthy controls.

The study was approved by the relevant Institutional Review Board and informed consent was provided. Adolescent inpatients were collectively assessed by doctoral-level clinical psychology students and/or trained clinical research assistants. Assessments were conducted independently and in private within the first 2 weeks following admission on the inpatient unit. Healthy controls were assessed by doctoral-level clinical psychology students and/or trained clinical research assistants in private assessment rooms at a local University or in schools where adolescents were recruited from.

Measures

Hypermentalizing questionnaire (HMZQ)

The HMZQ as described in Study 1 was administered to all participants.

The Childhood Interview for DSM-IV Borderline Personality Disorder (CI-BPD; Zanarini et al., 2003) is a semi-structured diagnostic interview for use with children and adolescents. The CI-BPD assesses the nine DSM-IV criteria of BPD, which were unchanged in Section II of the DSM-5. Each criterion has a set of corresponding prompts used by the interviewer to investigate that criterion, from which they rate with a score of 0 (absent), 1 (probably present), or 2 (definitely present). Adolescents who meet five or more criteria at the 2-level meet diagnostic criteria for BPD. Additionally, a Total Score can be used as a dimensional measure of BPD features, which is a sum of scores for each of the 9 criteria (maximum score of 18). Excellent psychometric properties for this measure have been demonstrated in a sample of inpatient adolescents (Sharp et al., 2012) as well as high concordance (94%) between parents and adolescents on BPD diagnoses based on use of

TABLE 4 Sample characteristics and HMZQ performance by group.

	BPD $n = 70$	Psychiatric Non-BPD $n = 136$	Healthy controls $n = 158$	Group comparisons
Demographics				
% Female	82.9	47.4	69.2	$\chi^2 = 14.09$, $df = 2{,}365$, $p = 0.001$
M Age (SD)	15.37 (1.52)	15.37 (1.30)	15.46 (1.28)	$F = 1.74$, $df = 2$, $p = 0.18$
%White/Not Hispanic	78.7	85.7	7.6	$\chi^2 = 222.75$, $df = 10{,}338$, $p < 0.001$
% Black	1.6	1.7	18.4	
% Hispanic/Not White	8.2	4.2	39.9	
% Multiracial or other	9.8	5	1.9	
Psychiatric comorbidity				
% Depressive disorder	72.9	65.4		$\chi^2 = 3.32$, $df = 1{,}193$, $p = 0.07$
% Bipolar disorder	11.4	3.8		$\chi^2 = 5.29$, $df = 1{,}193$, $p = 0.02$
% Eating disorder	20	6.6		$\chi^2 = 9.46$, $df = 1{,}193$, $p = 0.002$
% Externalizing disorder	51.4	33.1		$\chi^2 = 8.84$, $df = 1{,}193$, $p = 0.003$
% Anxiety disorder	71.4	59.6		$\chi^2 = 5.66$, $df = 1{,}193$, $p = 0.02$
HMZQ				
Total score (SD)	154.38 (42.07)	120.80 (56.41)	107.23 (51.66)	$F = 14.34$, $df = 2$, $p < 0.001$
Parent version (SD)	60.58 (20.24)	48.60 (19.44)	38.41 (20.29)	$F = 34.74$, $df = 2$, $p < 0.001$
Romantic partner (SD)	55.75 (18.94)	40.48 (23.22)	34.05 (22.87)	$F = 18.05$, $df = 2$, $p < 0.001$
Close friend (SD)	43.28 (22.35)	34.67 (21.81)	28.43 (18.01)	$F = 13.24$, $df = 2$, $p < 0.001$

40% of psychiatric controls, 25% of BPD patients, and 43% of healthy controls did not complete the romantic partner version of the HMZQ. Therefore, when calculating the total score on the HMZQ, group sizes were reduced due to the missing data for the romantic partner version of the measure to BPD (n = 55), psychiatric controls (n = 95) and healthy controls (n = 91). Psychiatric comorbidity determined using the Computerized Diagnostic Interview for DSM-IV (C-DISC; Shaffer et al., 2000) conducted with clients based on diagnostic criteria being met in the past year. The C-DISC was not completed in the healthy control sample.

this measure (Wall et al., 2017). Interrater reliability was evaluated on 13% ($n = 40$) of inpatient cases using Cohen's Kappa statistic. There was strong agreement between the original interviewer and an independent rater of the recorded interview for the final BPD diagnosis ($\kappa = 0.886$, $p < 0.001$).

Study 2 results

To determine the ability of the HMZQ to differentiate adolescents with BPD from those with other disorders and healthy controls, we compared the means of the various HMZQ scores across groups. Table 4 shows significant differences for all HMZQ scores with the largest effect size for the parent version of the measure. Tukey tests confirmed the superior performance of the parent version of the HMZQ by being the only HMZQ score that distinguished between all three subgroups. Specifically, comparisons of healthy controls vs. the BPD group (Tukey $= -27.66$; $p < 0.001$) and psychiatric controls vs. the BPD group (Tukey $= -10.91$; $p = 0.04$) showed significant differences for the parent version of the HMZQ. In contrast, only comparisons of healthy controls vs. the BPD group (Tukey $= -26.20$; $p < 0.001$), but not psychiatric controls vs. the BPD group (Tukey $= -12.04$; $p = 0.21$) showed significant differences for the romantic partner version of the HMZQ. Similarly, only comparisons of healthy controls vs. the BPD group (Tukey $= -14.76$; $p < 0.001$), but not psychiatric controls vs. the BPD group (Tukey $= -1.69$; $p = 0.98$) showed significant differences for the romantic partner version of the HMZQ. Taken together, only the total score and parent version of the HMZQ seems to be effective in distinguishing personality pathology from other psychopathology in youth. These results are depicted in Figure 1.

Discussion of study 2

The main goal of Study 2 was to assess the ability of the HMZQ to distinguish between psychiatric and non-psychiatric populations, and to investigate the specificity of HMZQ to borderline pathology. While our findings suggest that all versions of the HMZQ were good at distinguishing between healthy controls and borderline patients, only the total score and the parent version of the measure distinguished between adolescents with psychiatric disorders without BPD (psychiatric controls) and adolescents with BPD. This means that the romantic partner and best friend versions of the HMZQ are not sensitive to differences in groups with different psychopathology, but are most likely only sensitive to psychiatric severity in general. Given the theoretical roots of mentalizing in personality development and personality pathology, as discussed earlier, a measure that is sensitive to differences between general psychopathology and personality pathology would be considered more valid and fit for purpose. Given the attachment

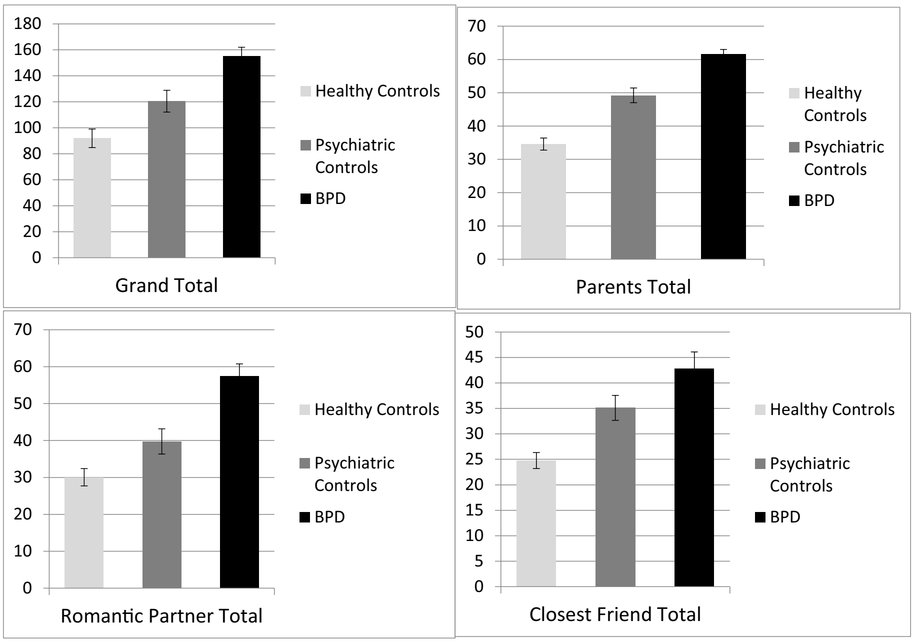


FIGURE 1
Group difference in HMZQ scores among adolescents meeting criteria for BPD, psychiatric controls, and healthy controls.

roots of the mentalizing construct as defined within Fonagy et al.'s (2002) model, it is perhaps not surprising that it is mentalizing in the context of the original attachment relationship—that is, the one with parents, that provide the best context for the assessment of hypermentalizing.

Overall discussion

Hypermentalizing is a relatively new construct which has resonated with clinicians and researchers who routinely work with borderline personality disorder (Bo et al., 2015). While clinicians have often recognized the tendency in their patients to hypermentalize, until recently there had not been an empirically grounded construct available to describe or assess this tendency. For instance, psychodynamic object relation therapies have used the term "projective identification" to refer to a process akin to hypermentalizing. Projective identification was introduced by Melanie Klein and is broadly defined as the process whereby in a close relationship (e.g., often an attachment relationship or a relationship between a therapist and patient), parts of the self may in unconscious fantasy be thought of as being forced into the other person (Casement, 1990). Projective identification serves an important defensive function for the individual. Specifically, feelings which cannot be consciously accessed are defensively projected into another person in order *to evoke* the thoughts or feelings projected (Jacobs, 2006). Hypermentalizing is also evident in cognitive-behavioral writing in the form of "mindreading errors" (e.g., Burns, 1980) defined as making negative interpretations even though there is no definite fact that convincingly support the conclusion; for example, one arbitrarily concludes that someone is reacting negatively to one and one does not bother to check this out. The advantage of the hypermentalizing construct is that it is tied to a particular experimental task (the Movie Assessment for Social Cognition; MASC) and is empirically and conceptually grounded in the social-cognitive literature with associated clinical, behavioral, cognitive, and neurobiological correlates (e.g., Badcock, 2011; Franzen et al., 2011; Gambin et al., 2015; Langdon and Brock, 2008; Langdon and Coltheart, 1999; Sharp and Vanwoerden, 2015). In our view, a particular attractive feature of the hypermentalizing construct is the carefully articulated formulation of its developmental roots in attachment theory which provide additional conceptual coherence.

The aim of the current study was fully operationalize the construct of hypermentalizing and to facilitate the use of this construct in clinical and research settings in adolescents and young adults by developing and evaluating a self-report measure of hypermentalizing. To this end, we conducted two studies. The first made use of a college sample to evaluate the purported factor structure of the HMZQ and assess the validity of the derived factor structure by using the MASC as external validity measure. The second study explored the clinical utility of the newly developed measure in a sample of adolescents comprised of three groups: borderline, non-borderline psychiatric controls and healthy controls. Across the two studies, results provided preliminary support for the use of the parent version of the HMZQ in particular. In Study 1, while the factor structure of the single trait and multi-trait/multi-method factor models were both supported, and both models accounted for incremental validity in predicting hypermentalizing on the MASC beyond the parent, close friend, and romantic partner total scores, the parent total score was the only one that demonstrated both convergent and discriminant validity—positively correlating with hypermentalizing, negatively correlating with correct mentalizing and not correlating with hypomentalizing or undermentalizing. Study 2, in a different sample, confirmed the superiority of the parent version of the HMZQ in that it was only the HMZQ version that distinguished not only BPD from healthy controls but also BPD from psychiatric controls. Results of Study 1 also suggested little evidence in support of using the subscale scores for the purported factors (an overconcern with the mental states of others; overinterpretation of others' mental states; inflexible certainty in own beliefs about others' mental states; acting impulsively on the assumed mental states of others; and second-guessing or over-interpretation of own mental states) and the use of the total score of the parent version is recommended.

That the parent version of the HMZQ outperforms the romantic partner and close friend versions of the HMZQ in its association with hypermentalizing while at the same time performing almost as well as the single trait factor score, means that the 26 items comprising the parent version is the most parsimonious and effective way of capturing the latent construct of hypermentalizing. As discussed, the construct of mentalizing and hypermentalizing are both theoretically and empirically grounded in attachment; it is therefore no surprise that individual differences in hypermentalizing in the context of parental relationships appears to be sufficient for explaining variance in relevant outcomes. The outcomes in this study included experimentally defined hypermentalizing and borderline personality disorder, so it is left to be seen if the same would be true for other outcomes. Even so, attachment to parents developmentally precedes attachment to peers and is seen as the basis on which attachment to peers and romantic partners are built.

The current study has several limitations. First, construct validity of the HMZQ is partly based on group comparisons of scores on this measure leading to the possibility of response bias based on group characteristics (Millsap, 2011); therefore, future research should test measurement invariance of responses between clinical and non-clinical groups to determine whether the questionnaire functions in the same way across groups. Similarly, measurement invariance over time should be tested to determine whether development has an influence in responses to this questionnaire given the aim for this measure to have utility across adolescence and adulthood. Second, there were significant differences in demographics across samples in Study 2—specifically females were over-represented in both samples and there were significant socio-demographic differences between samples. While gender differences are in line with previous findings of higher prevalence of borderline personality disorder among females in clinical samples (Sansone and Sanson, 2011), ethnoracial differences across samples can partly be accounted for by the socioeconomic differences between the sample recruited from the community and the inpatient sample, which due to cost typically serves families with high incomes. Therefore, findings must be replicated in clinical samples that are more representative in terms of socioeconomic status and ethnicity/race. Findings should also be replicated in samples of older adults, and individuals from different cultural background, since the findings of the current study are generalizable to adolescent and young adults in the US context only. Finally, as with any self-report measure, there is the potential for response processes that are unrelated to the construct under consideration. However, concerns over the validity of self-report of mentalizing are somewhat mitigated by the fact that over the last two decades, a significant number of studies have been published suggesting strong psychometric properties for self-report measures of mentalizing in adolescents (e.g., Sharp et al., 2009; Ha et al., 2013; Duval et al., 2018; Lund et al., 2022; Sharp et al., 2021).

Notwithstanding these limitations, the current study introduces a self-report measure of hypermentalizing and provide preliminary evidence in support of its further validation using other approaches. The advantage of such a measure is that it can be used in clinical settings to assess the level of hypermentalizing errors associated with general psychopathology as well as personality pathology. Our results clearly showed significant differences for the total and parent hypermentalizing scores between healthy controls, those with psychiatric disorder (but no personality pathology) and those with personality disorder, with the latter group evidencing the highest levels of hypermentalizing. A relatively short and easy-to-administer measure of hypermentalizing facilitates the identification of reducing hypermentalizing in individuals with all forms of psychopathology, thereby expanding the hypermentalizing construct beyond its most common current application in personality pathology research and practice.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by University of Houston Committee for the Protection of Human Subjects. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

CS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. CB: Formal analysis, Writing – original draft, Writing – review & editing. SV: Data curation, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abu-Akel, A. (2008). Theory of mind in autism, schizophrenia, and in-between. *Behav. Brain Sci.* 31:261. doi: 10.1017/S0140525X08004226

Abu-Akel, A., and Shamay-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia* 49, 2971–2984. doi: 10.1016/j.neuropsychologia.2011.07.012

Allen, J. G. (2003). Mentalizing. *Bull. Menninger Clin.* 67, 91–112. doi: 10.1521/bumc.67.2.91.23440

Badcock, C. (2011). *The Imprinted Brain*. London: Jessica Kingsley Publishers.

Baron-Cohen, S. (2000). "Theory of mind and autism: a fifteen year review," in *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience, 2nd Edn.*, eds. S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen (Oxford: Oxford University Press), 3–20. doi: 10.1093/oso/9780198524465.003.0001

Bateman, A. W., and Fonagy, P. (2012a). *Mentalization-Based Treatment of Borderline Personality Disorder*. New York, NY: Oxford University Press. doi: 10.1093/oxfordhb/9780199735013.013.0036

Bateman, A. W., and Fonagy, P. (2012b). *Handbook of Mentalizing in Mental Health Practice*. Washington, DC: American Psychiatric Publishing.

Belsky, D. W., Caspi, A., Arseneault, L., Bleidorn, W., Fonagy, P., Goodman, M., et al. (2012). Etiological features of borderline personality related characteristics in a birth cohort of 12-year-old children. *Dev. Psychopathol.* 24, 251–265. doi: 10.1017/S0954579411000812

Bo, S., Sharp, C., Fonagy, P., and Kongerslev, M. (2015). Hypermentalizing, attachment, and epistemic trust in adolescent BPD: clinical illustrations. *Personal Disord.* 8, 172–182. doi: 10.1037/per0000161

Brune, M., and Brune-Cohrs, U. (2006). Theory of mind - evolution, ontogeny, brain mechanisms and psychopathology. *Neurosci. Biobehav. Rev.* 30, 437–455. doi: 10.1016/j.neubiorev.2005.08.001

Burns, D. (1980). *Feeling Good: The New Mood Therapy*. New York, NY: Avon books.

Casement, P. (1990). *Further Learning from the Patient.* London: Routledge.

Crespi, B., and Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain. *Behav. Brain Sci.* 31:241. doi: 10.1017/S0140525X08004214

Dahl, R. E., Allen, N. B., Wilbrecht, L., and Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature* 554, 441–450. doi: 10.1038/nature25770

Duval, J., Ensink, K., Normandin, L., Sharp, C., and Fonagy, P. (2018). Measuring reflective functioning in adolescents: relations to personality disorders and psychological difficulties. *Adolesc. Psychiatry* 8, 5–20. doi: 10.2174/2210676608666180208161619

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., et al. (2006). Introducing MASC: a movie for the assessment of social cognition. *J. Autism Dev. Disord.* 36, 623–636. doi: 10.1007/s10803-006-0107-0

Ensink, K., Normandin, L., Target, M., Fonagy, P., Sabourin, S., and Berthelot, N. (2015). Mentalization in children and mothers in the context of trauma: an initial study of the validity of the Child Reflective Functioning Scale. *Br. J. Dev. Psychol.* 33, 203–217. doi: 10.1111/bjdp.12074

Fonagy, P. (1991). Thinking about thinking: some clinical and theoretical considerations in the treatment of a borderline patient. *Int. J. Psychoanal.* 72, 639–656.

Fonagy, P., Gergely, G., Jurist, E. L., and Target, M. (2002). *Affect Regulation, Mentalization, and the Development of Self*. New York, NY: Other Press.

Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y. W., Warren, F., Howard, S., et al. (2016). Development and validation of a self-report measure of mentalizing: the reflective functioning questionnaire. *PLoS ONE* 11:e0158678. doi: 10.1371/journal.pone.0158678

Fonagy, P., Steele, H., and Steele, M. (1991). Maternal representations of attachment during pregnancy predict the organization of infant-mother attachment at one year of age. *Child Dev.* 62, 891–905. doi: 10.2307/1131141

Franzen, N., Hagenhoff, M., Baer, N., Schmidt, A., Mier, D., Sammer, G., et al. (2011). Superior 'theory of mind' in borderline personality disorder: an analysis of interaction behavior in a virtual trust game. *Psychiatry Res.* 187, 224–233. doi: 10.1016/j.psychres.2010.11.012

Frick, C., Lang, S., Kotchoubey, B., Sieswerda, S., Dinu-Biringer, R., Berger, M., et al. (2012). Hypersensitivity in borderline personality disorder during mindreading. *PLoS ONE* 7:e41650. doi: 10.1371/journal.pone.0041650

Frith, C. D. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale, NJ: Erlbaum.

Fuligni, A. J., and Eccles, J. S. (1993). Perceived parent-child relationships and early adolescents orientation toward peers. *Dev. Psychol.* 29, 622–632. doi: 10.1037/0012-1649.29.4.622

Gambin, M., Gambin, T., and Sharp, C. (2015). Social cognition, psychopathological symptoms, and family functioning in a sample of inpatient adolescents using variable-centered and person-centered approaches. *J. Adolesc.* 45, 31–43. doi: 10.1016/j.adolescence.2015.08.010

Gorrese, A., and Ruggieri, R. (2012). Peer attachment: a meta-analytic reivew of gender and age differences and associateions with parent attachment. *J. Youth Adolesc.* 41, 650–672. doi: 10.1007/s10964-012-9759-6

Grant, P., Balser, M., Munk, A. J., Linder, J., and Hennig, J. (2014). A false-positive detection bias as a function of state and trait schizotypy in interaction with intelligence. *Front. Psychiatry* 5:135. doi: 10.3389/fpsyt.2014.00135

Ha, C., Sharp, C., Ensink, K., Fonagy, P., and Cirino, P. (2013). The measurement of reflective function in adolescents with and without borderline traits. *J. Adolesc.* 36, 1215–1223. doi: 10.1016/j.adolescence.2013.09.008

Hallquist, M. N., Hipwell, A. E., and Stepp, S. D. (2015). Poor self-control and harsh punishment in childhood prospectively predict borderline personality symptoms in adolescent girls. *J. Abnorm. Psychol.* 124, 549–564. doi: 10.1037/abn0000058

Happé, F. G. E., and Frith, U. (1996). Theory of mind and social impairment in children with conduct disorder. *Br. J. Dev. Psychol.* 14, 385–398. doi: 10.1111/j.2044-835X.1996.tb00713.x

Hezel, D. M., and McNally, R. J. (2014). Theory of mind impairments in social anxiety disorder. *Behav. Ther.* 45, 530–540. doi: 10.1016/j.beth.2014.02.010

Howes, O. D., and Kapur, S. (2009). The dopamine hypothesis of schizophrenia: version III - the final common pathway. *Schizophr. Bull.* 35, 549–562. doi: 10.1093/schbul/sbp006

Jacobs, M. (2006). *Psychodynamic Counselling in Action*. London: Sage.

Kim, S. (2015). The mind in the making: developmental and neurobiological origins of mentalizing. *Personal. Disord.* 6, 356–365. doi: 10.1037/per0000102

Langdon, R., and Brock, J. (2008). Hypo- or hyper-mentalizing: it all depends upon what one means by "mentalizing". *Behav. Brain Sci.* 31, 274–275. doi: 10.1017/S0140525X08004354

Langdon, R., and Coltheart, M. (1999). Mentalising, schizotypy, and schizophrenia. *Cognition* 71, 43–71. doi: 10.1016/S0010-0277(99)00018-9

Langdon, R., Coltheart, M., and Ward, P. B. (2006a). Empathetic perspective-taking is impaired in schizophrenia: evidence from a study of emotion attribution and theory of mind. *Cogn. Neuropsychiatry* 11, 133–155. doi: 10.1080/13546800444000218

Langdon, R., Corner, T., McLaren, J., Coltheart, M., and Ward, P. B. (2006b). Attentional orienting triggered by gaze in schizophrenia. *Neuropsychologia* 44, 417–429. doi: 10.1016/j.neuropsychologia.2005.05.020

Lecours, S., and Bouchard, M. A. (1997). Dimensions of mentalisation: outlining levels of psychic transformation. *Int. J. Psychoanal.* 78, 855–875.

Lund, S. H., Bo, S., Bach, B., Jørgensen, M. S., and Simonsen, E. (2022). Mentalizing in adolescents with and without prominent borderline features: validation of the reflective functioning questionnaire for youths (RFQY) and an investigation of the factor structure of hypo-and Hypermentalizing. *J. Pers. Assess.* 105, 475–486. doi: 10.1080/00223891.2022.2055474

Marty, P. (1991). *Mentalisation et Psychosomatique*. Paris: Laboratoire Delagrange.

Marty, P., and M'Uzan, D. (1963). La pensée opératoire. *Rev. Fr. Psychanal.* 27, 1345–1356.

McLaren, V., Hopwood, C., Gallagher, M., and Sharp, C. (2022). Hypermentalizing and borderline personality disorder: a meta-analytic review. *Am. J. Psychother.* 5, 21–31. doi: 10.1176/appi.psychotherapy.20210018

Meins, E. (1998). The effects of security of attachment and maternal attribution of meaning on children's linguistic acquisitional style. *Infant Behav. Dev.* 21, 237–252. doi: 10.1016/S0163-6383(98)90004-2

Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. New York, NY: Routledge. doi: 10.4324/9780203821961

Morton, J. (1989). *The Origins of Autism*. London: New Scientist.

O'Connor, T. G., and Hirsch, N. (1999). Intra-individual differences and relationship-specificity of mentalising in early adolescence. *Soc. Dev.* 8, 256–274. doi: 10.1111/1467-9507.00094

Preissler, S., Dziobek, I., Ritter, K., Heekeren, H. R., and Roepke, S. (2012). Social cognition in borderline personality disorder: evidence for disturbed recognition of the emotions, thoughts, and intentions of others. *Front. Behav. Neurosci.* 4:182. doi: 10.3389/fnbeh.2010.00182

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a 'theory of mind'? *Behav. Brain Sci.* 4, 515–526. doi: 10.1017/S0140525X00076512

Sansone, R. A., and Sanson, L. A. (2011). Gender patterns in borderline persoanlity disorder. *Innov. Clin. Neurosci.* 8, 16–20.

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., and Patton, G. C. (2018). The age of adolescence. *Lancet Child. Adolesc. Health* 2, 223–228. doi: 10.1016/S2352-4642(18)30022-1

Semerari, A., Carcione, A., Dimaggio, G., Nicolo, G., Pedone, R., and Procacci, M. (2005). Metarepresentative functions in borderline personality disorder. *J. Pers. Disord.* 19, 690–710. doi: 10.1521/pedi.2005.19.6.690

Semerari, A., Carcione, A., Dimaggio, G., Nicolo, G., and Procacci, M. (2007). Understanding minds: different functions and different disorders? The contribution of psychotherapy research. *Psychother. Res.* 17, 106–119. doi: 10.1080/10503300500536953

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M., and Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *J. Am. Acad. Child Adolsc.* 39, 28–38. doi: 10.1097/00004583-200001000-00014

Sharp, C. (2000). *Biased minds: theory of mind and emotional-behaviour disorders of middle childhood* (Unpublished Ph.D.). University of Cambridge, Cambridge, England.

Sharp, C. (2008). Theory of mind and conduct problems in children: deficits in reading the 'emotions of the eyes'. *Cogn. Emot.* 22, 1149–1158. doi: 10.1080/02699930701667586

Sharp, C. (2014). "The social-cognitive basis of borderline personality disorder: a theory of hypermentalizing," in *The Handbook of Borderline Personality Disorder in Children and Adolescents*, eds. C. Sharp and J. Tackett (New York, NY: Springer), 211–228. doi: 10.1007/978-1-4939-0591-1

Sharp, C., and Bevington, D. (2022). *Mentalizing in Psychotherapy: A Guide for Practitioners.* New York, NY: Guilford Publications, Inc.

Sharp, C., and Fonagy, P. (2008). The parent's capacity to treat the child as a psychological agent: constructs, measures and implications for developmental psychopathology. *Soc. Dev.* 17, 737–754. doi: 10.1111/j.1467-9507.2007.00457.x

Sharp, C., Ha, C., Michonski, J., Venta, A., and Carbone, C. (2012). Borderline personality disorder in adolescents: evidence in support of the Childhood Interview for DSM-IV Borderline Personality Disorder in a sample of adolescent inpatients. *Compr. Psychiatry* 53, 765–774. doi: 10.1016/j.comppsych.2011.12.003

Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A. B., Sturek, J., et al. (2011). Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *J. Am. Acad. Child Adolesc. Psychiatry* 50, 563–573. doi: 10.1016/j.jaac.2011.01.017

Sharp, C., Steinberg, L., McLaren, V., Weir, S., Ha, C., and Fonagy, P. (2021). Refinement of the reflective function questionnaire for youth (RFQY) using item response theory. *Assessment* 29(6):1204–1215. doi: 10.1177/10731911211003971

Sharp, C., and Vanwoerden, S. (2015). Hypermentalizing in borderline personality disorder: a model and data. *J. Infant Child Adolesc. Psychother.* 14, 33–45. doi: 10.1080/15289168.2015.1004890

Sharp, C., Vanwoerden, S., Van Baardewijk, Y., Tackett, J. L., and Stegge, H. (2015a). Callous-unemotional traits are associated with deficits in recognizing complex emotions in preadolescent children. *J. Pers. Disord.* 29, 347–359. doi: 10.1521/pedi_2014_28_161

Sharp, C., and Venta, A. (2012). "Mentalizing problems in children and adolescents," in *Minding the Child: Mentalization-Based Interventions with Children, Young People and their Families*, eds. N. Midgley and I. Vrouva (London: Routledge), 35–53.

Sharp, C., Venta, A., Ha, C., Schramm, A., Reddy, R., and Newlin, E. (2015b). First evidence for the link between attachment, social cognition and borderline features in adolescents. *Comprehens. Psychiatry*. 64, 4–11. doi: 10.1016/j.comppsych.2015.07.008

Sharp, C., Williams, L. W. W., Ha, C., Baumgardner, J., Michonski, J., Seals, R., et al. (2009). The development of a mentalization-based outcomes and research protocol for an adolescent in-patient unit. *Bull. Menninger Clinic.* 73, 311–338. doi: 10.1521/bumc.2009.73.4.311

Steinberg, L., Dahl, R. E., Keating, D., Kupfer, D. J., Masten, A., and Pine, D. (2006). "The study of developmental psychopathology in adolescence: integrating affective neuorsicence with the study of context," in *Developmental Psychopathology*, Vol. 2, eds. D. Cicchetti and D. Cohen (New York, NY: Wiley), 710–741. doi: 10.1002/9780470939390.ch18

Symons, D. K., and Clark, S. E. (2000). A longitudinal sutdy of mother-child relationships and thoery of mind in the preschool period. *Soc. Dev.* 9, 2–23. doi: 10.1111/1467-9507.00108

Wall, K., Sharp, C., Ahmed, Y., Goodman, M., and Zanarini, M. C. (2017). Parent-adolescent concordance on the Revised Diagnostic Interview for Borderlines (DIB-R) and the Childhood Interview for Borderline Personality Disorder (CI-BPD). *Personal. Ment. Health* 11, 179–188.

Washburn, D., Wilson, G., Roes, M., Rnic, K., and Harkness, K. L. (2016). Theory of mind in social anxiety disorder, depression, and comorbid conditions. *J. Anxiety Disord.* 37, 71–77. doi: 10.1016/j.janxdis.2015.11.004

Zanarini, M. C., Vujanovic, A. A., Parachini, E. A., Boulanger, J. L., Frankenburg, F. R., and Hennen, J. (2003). Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J. Pers. Disord.* 17, 233–242. doi: 10.1521/pedi.17.3.233.22147

TYPE Original Research
PUBLISHED 17 July 2025
DOI 10.3389/fpsyg.2025.1549378

OPEN ACCESS

EDITED BY
Antonella Marchetti,
Catholic University of the Sacred Heart, Italy

REVIEWED BY
Elisabetta Lombardi,
eCampus University, Italy
Federica Bianco,
University of Bergamo, Italy

*CORRESPONDENCE
Daniel M. Bernstein
✉ daniel.bernstein@kpu.ca

RECEIVED 21 December 2024
ACCEPTED 16 June 2025
PUBLISHED 17 July 2025

CITATION
Erceg HG, Dhillon RS, Derksen DG, Mah EY and Bernstein DM (2025) A longitudinal study of theory of mind across the lifespan. *Front. Psychol.* 16:1549378.
doi: 10.3389/fpsyg.2025.1549378

A longitudinal study of theory of mind across the lifespan

Hanna G. Erceg[1,2], Ruby S. Dhillon[2], Daniel G. Derksen[2], Eric Y. Mah[3] and Daniel M. Bernstein[1,2]*

[1]Department of Psychology, Kwantlen Polytechnic University, Surrey, BC, Canada, [2]Department of Psychology, Simon Fraser University, Burnaby, BC, Canada, [3]Department of Psychology, University of Victoria, Victoria, BC, Canada

Introduction: Theory of Mind (ToM) is essential for social interactions. However, gaps remain in our knowledge of when ToM abilities develop and change, particularly from adolescence to older adulthood.

Methods: We used data from an ongoing longitudinal study to examine ToM abilities across three time points in participants aged 3 years and older. Testing waves occurred over multiple years. Cognitive ToM was assessed using the Sandbox task ($N = 187$; age range = 3–80 years), and affective ToM was assessed using the Reading the Mind in the Eyes Task (RMET; $N = 121$; age range = 6–80 years). Data were analyzed using mixed-design ANOVAs to examine interactions between Age Group and Time Point.

Results: Children aged 6–9 years exhibited significantly lower ToM abilities compared to adults. However, beyond childhood, both cognitive and affective ToM remained relatively stable across the lifespan.

Discussion: Our study illuminates critical periods of ToM development. Moreover, our study highlights the importance of using measures that capture subtle changes across the lifespan.

KEYWORDS

theory of mind, longitudinal analysis, lifespan, Sandbox task, reading the mind in the eyes, social cognition, cognitive and affective ToM, false-belief understanding

Introduction

A longitudinal study of theory of mind across the lifespan

Theory of Mind (ToM) is the ability to attribute mental states to oneself and others. Specifically, ToM is the ability to understand and reason about beliefs, desires, thoughts, intentions, and feelings (Premack and Woodruff, 1978; Wimmer and Perner, 1983). ToM plays a crucial role in everyday social interactions. However, despite over four decades of research, gaps remain in our knowledge of when ToM abilities develop and change, particularly across the adolescent to older adult lifespan (Derksen et al., 2018). While developmental patterns of ToM have been explored, much of our knowledge comes from cross-sectional studies. There is, however, longitudinal research focusing on children (see Wellman, 2014). Our study utilized a longitudinal design to explore ToM abilities across the child to older adult lifespan to provide a more comprehensive understanding of when these abilities develop and change. Our results highlight critical periods of ToM development.

Developmental patterns

Research has explored the developmental patterns of ToM abilities. Cross-sectional studies suggest that ToM abilities improve from preschool age to adolescence, stabilize from adolescence through adulthood, and then decline in older adulthood (Cornaggia et al., 2024; Dumontheil et al., 2010; Henry et al., 2013; Kong et al., 2025; Miller, 2022; Tousignant et al., 2017; Wellman et al., 2001). Notably, both cross-sectional and longitudinal research has primarily focused on children aged 3-13. These studies suggest that ToM mastery follows a predictable development of related skills in the following order: (a) diverse desires (understanding that different people can want different things), (b) diverse beliefs (understanding that opinions can differ), (c) knowledge access (not seeing = ignorance), (d) false belief understanding, (e) hidden emotion (people can conceal their true feelings behind false facial expressions), and (f) sarcasm (Peterson and Wellman, 2019; Wellman et al., 2011; Wellman and Liu, 2004)[1]. Research beyond middle childhood, especially longitudinal studies, is relatively sparse (see Derksen et al., 2018). As a result, there is a heavy reliance on cross-sectional designs, which limits the conclusions that can be drawn from the existing literature.

It remains unclear whether the observed developmental differences across the adolescent to older adult lifespan reflect true age-related changes or are merely the result of different task demands across the various measures used to assess ToM in different age groups. It is also possible that the observed differences reflect the development of the various skills necessary to complete ToM tasks (e.g., executive function, working memory) that vary across ToM tasks of different complexities. To address these concerns, longitudinal research using a single task to assess ToM across age groups is needed. Notably, discrete measures of ToM, such as the categorical change-in-location task known as the Sally-Anne task (Wimmer and Perner, 1983), are more common than continuous measures (e.g., reaction time, eye-tracking, mouse trajectory; Apperly et al., 2011; Keysar et al., 2003; O'Connor et al., 2024). Discrete measures may oversimplify the developmental trajectory of ToM, potentially exaggerating differences between age groups. For instance, a pass/fail coding system could make developmental changes appear more pronounced by masking subtle, continuous development. Thus, developmental patterns in ToM might be more subtle than previously believed.

To address concerns about the use of different ToM tasks across age groups and the limitations of discrete measures in capturing subtle developmental changes, Sommerville et al. (2013) developed the Sandbox task to measure ToM as a continuous (rather than categorical) variable. The Sandbox task is a modified change-of-location task appropriate for measuring false-belief understanding in preschoolers through older adults and also great apes (Lurz et al., 2022; c.f., Haskaraca et al., 2023; Samuel et al., 2018). Research using the Sandbox task reveals differing developmental patterns from those found using discrete measures in different age groups. While the existing literature shows striking developmental differences in ToM abilities across the child to older adult lifespan, cross-sectional work using the Sandbox task reveals that ToM abilities remain relatively stable from preschool to older age (Bernstein, 2021).

Components of theory of mind

Utilizing a single measure across age groups can address some limitations of past research, which has used various measures to assess ToM. However, understanding the developmental patterns of the distinct components of ToM is equally important. Notably, ToM consists of two main components: cognitive and affective (Shamay-Tsoory and Aharon-Peretz, 2007; see also Meinhardt-Injac et al., 2020). Cognitive ToM refers to the ability to understand the beliefs, intentions, and desires of oneself and others, while affective ToM refers to the ability to recognize and understand the emotions and feelings of others. Fewer studies have explored affective ToM than cognitive ToM (Mahy, 2018).

One way to measure affective ToM is with the Reading the Mind in the Eyes Task (RMET; Baron-Cohen et al., 2001). In the RMET, participants view pictures of eyes and indicate the matching emotion from a list. The RMET presents pictures of people's complex emotions in social situations. This is considered an advanced ToM ability (c.f., Higgins et al., 2024; Oakley et al., 2016) because a relevant social context must be referenced from memory to understand the emotion (Baron-Cohen et al., 2001, 1997a,b).

Affective ToM tends to decline earlier than cognitive ToM in older adults, highlighting a need for research to distinguish between the two components (Raimo et al., 2022; c.f., Bottiroli et al., 2016). One possible explanation for affective ToM declining earlier might relate to social interaction. Social interaction influences the development of ToM abilities. Increased social interactions enhance these abilities, presumably by providing more opportunities to practice inferring others' mental states (Yu and Wellman, 2023). Therefore, it is plausible that the observed decline in affective ToM around age 60 is related to the decrease in social interactions older adults experience during this life stage. However, a more likely explanation might be that declines in affective ToM are due to age-related declines in episodic memory, which is a specific cognitive ability. As noted earlier, the RMET requires participants to match expressions around the eyes to stored examples of relevant context from past experiences (Baron-Cohen et al., 2001). As episodic memory declines with age (Levine et al., 2002; Rönnlund et al., 2005), older adults may struggle to retrieve these episodes, leading to poorer affective ToM. Alternatively, cognitive ToM may decline later because of age-related declines in more general cognitive abilities. This understanding is supported by work revealing that executive function mediates age-related declines in cognitive ToM (Charlton et al., 2009; Phillips et al., 2011). Moreover, some work suggests that such age-related declines in cognitive ToM are due to age-related changes in executive functioning and not merely to declines in ToM competence (Cho and Cohen, 2019). However, this view is debated, and other work suggests cognitive ToM declines due to factors beyond task demands or general cognitive abilities, reflecting a decline in ToM competence itself (Bernstein et al., 2011; Bloom and German, 2000). Overall, the literature to date suggests future work would

1 This pattern varies slightly across cultures. Individuals from collectivist cultures are more likely to understand knowledge access before diverse beliefs, compared to those from individualist cultures (Shahaeian et al., 2011; Wellman et al., 2006).

benefit from measuring both cognitive and affective ToM abilities as distinct constructs.

The present study

To expand the existing literature on ToM, there is a need for longitudinal research that uses a single measure of ToM across age groups and distinguishes between cognitive and affective components. While longitudinal research on ToM exists, it has largely focused on preschool-aged children and adolescents (see Derksen et al., 2018). To our knowledge, there is currently no longitudinal work on ToM in adults.

This study seeks to advance our understanding of the developmental trajectory of ToM across the lifespan. We conducted a longitudinal analysis to explore age-related changes in ToM ability. We included separate measures of cognitive and affective ToM and used the same tasks across different age groups, spanning preschool to older adulthood. Based on previous literature, we hypothesized that: (1) Cognitive ToM would remain relatively stable from preschool to adulthood, with modest declines in older adulthood; (2) Affective ToM would remain relatively stable from childhood to adulthood, with modest declines in older adulthood emerging earlier compared to cognitive ToM.

Materials and methods

This research was conducted using data collected from an ongoing longitudinal study at a mid-sized University in Western Canada. Recruitment for the study started in 2015. Participants were recruited through various strategies. Children were recruited through local schools and community events, and older adults were recruited through community centers and independent living facilities. Undergraduate students were primarily recruited from the university's subject pool.

Participants completed a battery of measures, including the Sandbox task (Sommerville et al., 2013) and the RMET (Olderbak et al., 2015). To allow for within-subject comparisons over time, we limited our analyses to participants who had completed three waves of testing for each task. This decision preserved the integrity of the mixed-design ANOVAs, which require repeated measures across all included time points. For participants with more than three waves of data, we included their first three waves in the analysis.

A total of 696 participants completed at least one wave of testing for the Sandbox task, and 588 participants completed at least one wave of testing for the RMET. However, only a subset of participants returned and completed additional waves of testing for each task. As a result, a total of 187 participants completed three waves of testing for the Sandbox task (65.2% female, 34.8% male; mean age = 28.4 years, $SD = 24.9$, range = 3.06–80.2 years) and 121 participants completed three waves of testing for the RMET (67.8% female, 32.2% male; mean age = 31.1 years, $SD = 24.3$, range = 6.1–80.2 years). The average delay between time points for the Sandbox sample was 2.17 years ($SD = 1.73$), and 2.3 years ($SD = 1.67$) for the RMET sample.

Participants were grouped into age categories to reflect developmental stages (see Bernstein, 2021). Age groups for the Sandbox task included: 3–5 years ($N = 31$), 6–9 years ($N = 40$), 10–17 years ($N = 32$), 18–64 years ($N = 58$), 65+ years ($N = 26$). Age groups for the RMET included: 6–9 years ($N = 36$), 10–17 years ($N = 25$), 18–64 years ($N = 45$), 65+ years ($N = 15$). There were some differences in the average delay between time points across age groups (e.g., 65+ years had longer delays between waves than some other age groups, and the delay between Time Points 2–3 was on average longer than the delay between Time Points 1–2; see Supplementary material 1b and 2b for more details).

Measures

Sandbox task

The Sandbox Task is a modified change-of-location task used to measure cognitive ToM. Specifically, the Sandbox Task measures false-belief understanding as a continuous variable (Sommerville et al., 2013). Participants hear four short stories. In each story, an experimenter buries an object in a large box filled with Styrofoam at an initial location in the protagonist's view (L1). The protagonist then leaves. While the protagonist is absent, a second character digs up the item and moves it to a new location (L2), once again burying the item so that it is out of view from the protagonist. Participants then complete a 20-s visual search filler task before answering a false-belief and/or a memory-control question. In false belief questions, participants indicate where in the sandbox the protagonist would look for the item upon returning (requiring them to adopt the protagonist's perspective). In memory-control questions, participants are asked to recall where the item was initially placed (L1). In both cases participants respond by pointing to a location in the Sandbox. Experimenters record the response using a tape measure along the Sandbox's inside seam (visible only to the experimenter). We administered the task to participants 3 years and older. Prior to 2018, participants answered only one test question at the end of each trial: either a false-belief question or a memory-control question. Starting in 2018, participants answered both a false-belief and memory-control question for each story, doubling the amount of data collected. The majority (63%) of testing instances were collected using the single-question version of the task.

Bias scores were calculated separately for false-belief and memory-control trials. For each test question, responses toward the incorrect location denoted positive bias, and responses away from the incorrect location denoted negative bias (see Figure 1). An egocentric bias score was then calculated by subtracting the memory-control bias score from the false-belief bias score. Higher egocentric bias scores indicate greater difficulty suppressing one's own knowledge of the true location (L1) when reasoning about others' perspectives. Thus, a higher egocentric bias score reflects poorer false-belief reasoning, a key aspect of cognitive ToM.

Reading the mind in the eyes task (RMET)

We administered the RMET to assess affective ToM in participants ages 6 years and older (Baron-Cohen et al., 2001).

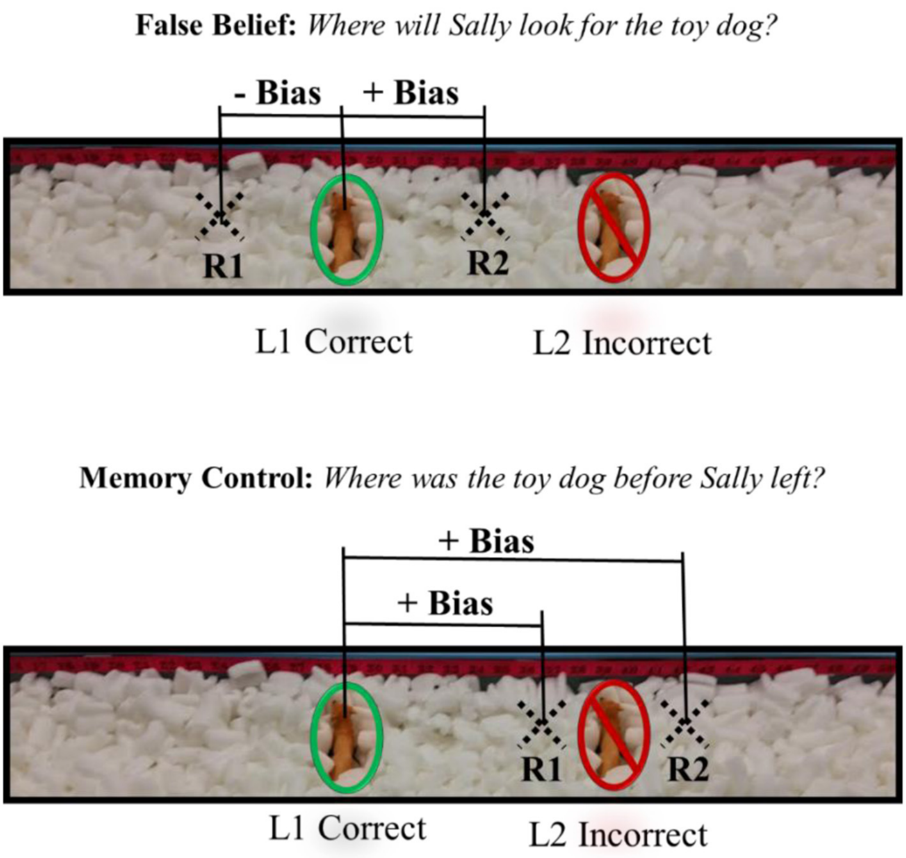


FIGURE 1
Bias calculation in the Sandbox task. L1 refers to the original location of the hidden object. L2 refers to the new location of the object after it was moved. R1 and R2 are examples of possible responses (i.e., Response 1 and Response 2). In all cases the object was not visible to the participant during their responses to test questions. In the False Belief example, a response at R1 would produce a negative bias because the response moves away from the incorrect location. A response at R2 would produce a positive bias because the response moves toward the incorrect location. In the Memory Control example, both responses would produce positive bias, however, a response at R2 would produce a larger positive bias as the response is even further from the correct location.

For adults, we initially used a shortened 12-item version of the RMET, but starting in 2019, we adjusted this to a 10-item version, which demonstrated better reliability compared to the original 36-item version (Olderbak et al., 2015). In our analysis sample (restricted to the first three completed waves of data for participants who had at least three waves of data), only one testing instance used the 10-item version. Children aged 6 to 17 years consistently received a 12-item version adapted with age-appropriate vocabulary. Preschoolers did not complete the RMET. Participants viewed a series of grayscale photographs depicting only the eye regions of various individuals. After each photograph appeared, participants tried to identify the emotional state that best represented the individual by choosing among four options. To ensure comprehension, participants received a list of definitions for each emotional state option. There was no time limit on the task. To account for the use of different task versions (i.e., 10 or 12 items), we calculated the proportion of correct responses. For adult data collected before January 2019, scores were divided by 12; adult data collected after January 2019 were divided by 10. Higher proportions indicate better affective ToM ability. Missing responses were treated at the trial level as incorrect.

Results

Prediction 1: cognitive ToM would remain relatively stable from preschool to adulthood, with modest declines in older adulthood

To compare differences in cognitive ToM ability across time points within different age groups, we conducted a 3 [Time Point: Time 1, Time 2, Time 3 (within)] × 5 [Age Group: 3–5 years, 6–9 years, 10–17 years, 18–64 years, 65+ years (between)] mixed-design ANOVA with egocentric bias as the dependent variable[2]. Assumptions of normality, homogeneity, and sphericity were violated; to avoid an inflated Type 1 error rate, a Greenhouse-Geisser correction was applied to adjust the degrees of freedom (Myers et al., 2010). There was a significant main effect of Age Group, $F_{(4,182)} = 2.959$, $p = 0.021$, $\eta^2G = 0.021$ (see Table 1

2 Analyses controlling for delays between time points and the number of questions asked are presented in the Supplementary materials.

10.3389/fpsyg.2025.1549378

for descriptive statistics). The main effect of Time Point was not significant, $F_{(1.62,295.29)} = 1.932$, $p = 0.155$, $\eta^2G = 0.007$. Thus, egocentric bias scores did not significantly change across Time Points. Further, the interaction between Age Group and Time Point was not significant, $F_{(6.49,295.29)} = 1.214$, $p = 0.297$, $\eta^2G = 0.018$ (see Figure 2 for visual representations of these trends). Thus, the pattern of differences in egocentric bias scores between Age Groups remained consistent across Time Points.

Pairwise comparisons were conducted to investigate significant differences between Age Groups (see Table 2). A Bonferroni correction was applied to account for multiple comparisons ($\alpha = 0.004$). After adjustment, only comparisons between the 6–9 years and 18–64 years groups were statistically significant, ($p = 0.0021$, $d = 0.395$) [3]. Participants aged 6–9 years exhibited higher egocentric bias scores than those aged 18–64 years, suggesting poorer cognitive ToM abilities in the younger age group.

Given assumption violations in the standard ANOVA, we conducted a trimmed ANOVA using 20% trimmed means to account for outliers and non-normality. The trimmed ANOVA indicated the main effect of Age Group was no longer significant, $F_{(4,46.5834)} = 1.5016$, $p = 0.2171$. These results suggest that the significant main effect of Age Group observed in the standard ANOVA may have been influenced by outliers. However, there was a significant main effect of Time Point, $F_{(2,54.9906)} = 3.3467$, $p = 0.0425$. Thus, there was a significant difference in egocentric bias scores across the three time points, suggesting that the participants' scores varied and did not remain stable throughout testing waves. The interaction between Age Group and Time Point remained non-significant, $F_{(8,52.6446)} = 0.4186$, $p = 0.9047$.

Overall, these results suggest that cognitive ToM abilities remained relatively stable within participants over the three time points, as evidenced by the lack of a significant interaction between Age Group and Time Point in both the standard ANOVA and the trimmed ANOVA.

TABLE 1 Means and standard deviations summary statistics for egocentric bias scores in the Sandbox Task for age groups across time points.

Age group	*N*	Time 1	Time 2	Time 3
3–5 years	31	3.18 *(9.74)*	0.377 *(12.1)*	3.66 *(5.86)*
6–9 years	40	4.62 *(7.20)*	2.29 *(9.66)*	2.23 *(5.67)*
10–17 years	32	2.54 *(5.33)*	3.10 *(7.12)*	0.332 *(2.71)*
18–64 years	58	0.753 *(5.66)*	0.812 *(4.91)*	0.477 *(2.11)*
65+ years	26	1.92 *(5.60)*	−0.144 *(2.59)*	1.45 *(3.45)*

Data are presented as means; italicized values are standard deviations (SD).

3 See Table 2 for all pairwise comparisons, including those significant before Bonferroni adjustment.

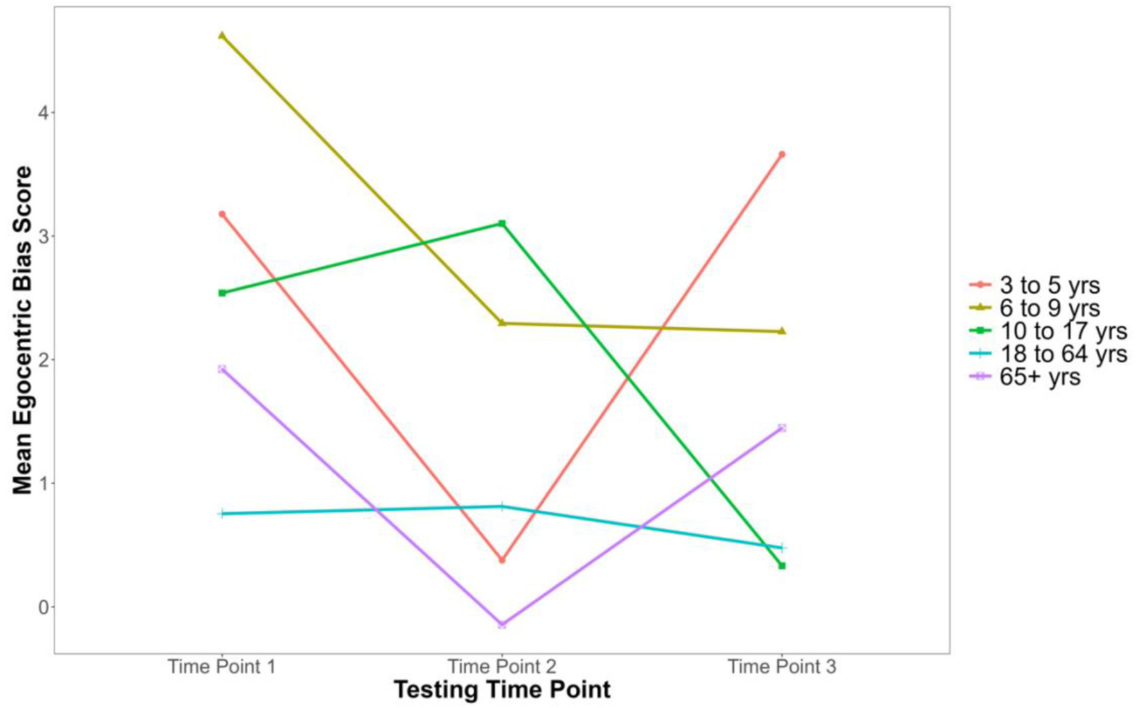


FIGURE 2
Mean egocentric bias scores in the Sandbox Task across time points for age groups. This graph is unscaled to emphasize differences between age groups and time points.

TABLE 2 Pairwise comparisons for egocentric bias scores in the Sandbox Task.

Age groups comparison	N1	N2	p (unadjusted)	Significance level
3–5 vs. 6–9	31	40	0.375	ns
3–5 vs. 10–17	31	32	0.709	ns
3–5 vs. 18–64	31	58	0.11	ns
3–5 vs. 65+	31	26	0.487	ns
6–9 vs. 10–17	40	32	0.197	ns
6–9 vs. 18–64	40	58	**0.0061**	***
6–9 vs. 65+	40	26	0.116	ns
10–17 vs. 18–64	32	58	0.233	ns
10–17 vs. 65+	32	26	0.731	ns
18–64 vs. 65+	58	26	0.466	ns

N1 and N2 represent the collapsed sample sizes for the first and second age groups in each comparison. ns, Not significant. ***p < 0.004 (Bonferroni threshold). Results significant at the Bonferroni threshold are bolded.

TABLE 3 Means and standard deviations summary statistics for correct recognition response scores in the RMET for age groups across time points.

Age group	N	Time 1	Time 2	Time 3
6–9 years	36	56.9 *(14.0)*	64.6 *(15.2)*	73.6 *(12.5)*
10–17 years	25	66.0 *(13.4)*	74.7 *(15.1)*	68.0 *(14.8)*
18–64 years	45	75.0 *(16.9)*	74.1 *(15.6)*	70.2 *(16.1)*
65+ years	15	73.3 *(13.4)*	75.6 *(14.9)*	62.2 *(15.1)*

Data are presented as means; italicized values are standard deviations (SD).

Prediction 2: affective ToM would remain relatively stable from preschool to adulthood, with modest declines in older adulthood emerging earlier compared to cognitive ToM

To compare differences in affective ToM ability across time points within different age groups, we conducted a 3 [Time Point: Time 1, Time 2, Time 3 (within)] × 4 [Age Group: 6–9 years, 10–17 years, 18–64 years, 65+ years (between)] mixed-design ANOVA with the correct recognition response score on the Eyes task as the dependent variable. Assumptions of normality, homogeneity, and sphericity were met. There was a significant main effect of Age Group, $F_{(3,117)} = 3.431$, $p = 0.019$, $\eta^2G = 0.047$. There was also a significant main effect of Time Point, $F_{(2,234)} = 3.875$, $p = 0.022$, $\eta^2G = 0.014$. Thus, affective ToM differed significantly both between Age Groups and across Time Points (see Table 3 for descriptive statistics). Further, the interaction between Age Group and Time Point was significant, $F_{(6,234)} = 8.112$, $p < 0.001$, $\eta^2G = 0.083$ (see Figure 3 for visual representations of these trends). Thus, the effect of Time Point on affective ToM varied across Age Groups.

Pairwise comparisons were conducted to investigate the interaction (see Table 4). A Bonferroni correction was applied to account for multiple comparisons ($\alpha = 0.004$). After adjustment, statistically significant differences were observed at Time 1 between the 6–9 years and 18–64 years Age Groups, ($p < 0.001$, $d = 1.154$), and between the 6–9 years and 65+ years Age Groups, ($p < 0.001$, $d = 1.186$). These results indicate that participants aged 6–9 years exhibited lower correct recognition response scores compared to participants aged 18–64 and 65+ years, suggesting poorer affective ToM abilities in the younger group. No significant differences were found at Time 2 or 3.

Discussion

Using a longitudinal design, we explored the developmental patterns of cognitive and affective ToM across the lifespan. We extended previous work by employing continuous and consistent measures of ToM from preschool age through older adulthood. We predicted that ToM would remain relatively stable into adulthood, with modest declines in older adulthood. Our results partially supported these predictions, and revealed similarities in the developmental trajectories of cognitive and affective ToM.

As predicted, cognitive ToM remained relatively stable across the lifespan, as indicated by the non-significant interaction between Age Group and Time Point in the standard and trimmed ANOVAs. However, there was a significant main effect of Age Group in the standard ANOVA. Specifically, participants aged 6–9 years exhibited significantly lower cognitive ToM (higher egocentric bias scores) compared to the 18–64 years group. This suggests there are developmental improvements in cognitive ToM during childhood, followed by stability across adulthood. This finding largely aligns with developmental patterns reported in prior research using continuous measures. For example, Bernstein (2021) observed that cognitive ToM abilities (i.e., false-belief reasoning) remained relatively stable from preschool to older adulthood. Moreover, Bernstein et al. (2017) observed relative stability in ToM abilities across most of the lifespan, with modest declines emerging in older adulthood. Contrary to our prediction, there was no evidence of cognitive ToM decline in older adulthood. Pairwise comparisons revealed no significant differences in egocentric bias scores between younger adults (18–64 years) and older adults (65+ years). Thus, cognitive ToM appears to remain relatively stable across much of the lifespan past childhood, at least within the timeframe measured in this study using the Sandbox task. Notably, these results should be interpreted with caution given that the assumptions of the standard ANOVA were violated.

Similarly, our prediction for affective ToM was only partially supported, with its developmental patterns revealing similarities to cognitive ToM. Results revealed a significant interaction between Age Group and Time Point. Specifically, at Time 1, participants aged 6–9 years demonstrated significantly lower affective ToM (lower correct recognition responses) compared to the 18–64 and 65+ age groups. Notably, the significant difference for the 65+ age group was not found for cognitive ToM, demonstrating similar yet distinct developmental trajectories. However, no other pairwise comparisons were statistically significant at either time point, suggesting that, beyond childhood, affective ToM remains relatively stable. Indeed, contrary to our prediction, there was no evidence of affective ToM decline in older adulthood. Thus, neither

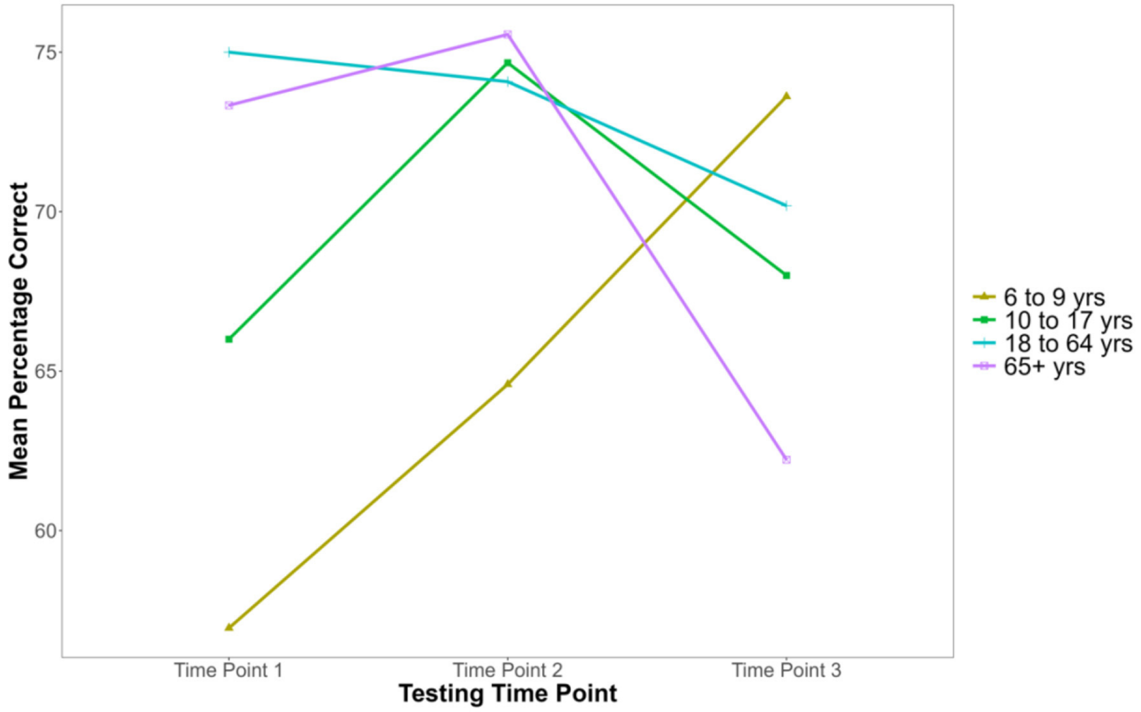


FIGURE 3
Mean percentage correct recognition response scores in the RMET across time points for age groups. This graph is unscaled to emphasize differences between age groups and time points.

ToM component demonstrated declines in older adulthood. Taken together, these results suggest that cognitive and affective components of ToM remain largely stable across the lifespan, with developmental changes occurring between childhood and adulthood.

Our findings align with prior research using continuous measures, which suggest that developmental patterns of ToM are relatively subtle (e.g., Bernstein et al., 2011; Dumontheil et al., 2010; Henry et al., 2013). This supports the view that continuous measures may better capture nuanced age-related changes compared to traditional discrete pass/fail measures. However, we acknowledge that this claim is premature, and that further work is needed to explore the differences between discrete and continuous measures of ToM. To address this, future studies could administer both continuous and discrete measures in a within-subject design to directly compare their ability to capture subtle changes in ToM across the lifespan. Using a wider variety of continuous (i.e., implicit) measures, such as reaction time (Kikuno et al., 2007), eye-tracking (Keysar et al., 2003), and mouse-tracking (van der Wel et al., 2014), would improve our understanding of nuanced age-related changes. Additionally, since continuous measures of ToM are less abundant, researchers might administer a battery of tasks that could then be combined into a continuous measure of ToM. This approach would also address the limitation of relying on a single discrete task. Overall, our findings emphasize the need to consider cognitive and affective ToM as distinct constructs that share similar developmental trajectories.

Beyond concerns related to measurement format (i.e., discrete vs. continuous), another important factor that may influence developmental patterns of ToM is task modality. Bottiroli et al. (2016) highlighted that differences in age-related ToM performance across studies may depend on whether the task relies on verbal or visual processing. Specifically, they proposed that abilities measured with verbal tasks (e.g., the Sandbox Task) remain relatively stable across the lifespan, as these tasks draw on verbal skills such as comprehension and social reasoning, which are relatively preserved with age. In contrast, performance on visual tasks (e.g., RMET) tends to decline earlier, as aging interferes with the ability to recognize emotions from facial expressions. Supporting this interpretation, Raimo et al. (2022) found that age-related declines in affective ToM were specific to tasks relying on visual modalities, whereas performance on verbal tasks remained relatively preserved. Future studies should carefully consider how task modality may affect observed developmental trajectories of ToM across the lifespan for both cognitive and affective ToM.

Methodological considerations and limitations

We used a single measure each for cognitive and affective ToM across age groups to address a key concern in the ToM literature regarding whether previously reported developmental differences reflect true age-related changes. However, this approach is also a

TABLE 4 Pairwise comparisons for correct recognition response scores in the RMET across time points.

Time point	Age groups comparison	*N1*	*N2*	*p* (unadjusted)	Significance level
1	6–9 vs. 10–17	36	25	0.0218	*
	6–9 vs. 18–64	36	45	**0.000000361**	****
	6–9 vs. 65+	36	15	**0.00053**	****
	10–17 vs. 18–64	25	45	0.0175	*
	10–17 vs. 65+	25	15	0.136	ns
	18–64 vs. 65+	45	15	0.709	ns
2	6–9 vs. 10–17	36	25	0.0128	*
	6–9 vs. 18–64	36	45	0.00649	**
	6–9 vs. 65+	36	15	0.0214	*
	10–17 vs. 18–64	25	45	0.877	ns
	10–17 vs. 65+	25	15	0.859	ns
	18–64 vs. 65+	45	15	0.746	ns
3	6–9 vs. 10–17	36	25	0.146	ns
	6–9 vs. 18–64	36	34	0.3	ns
	6–9 vs. 65+	36	15	0.0132	*
	10–17 vs. 18–64	25	45	0.553	ns
	10–17 vs. 65+	25	15	0.232	ns
	18–64 vs. 65+	45	15	0.0722	ns

N1 and N2 represent the sample sizes for the first and second age groups in each comparison. ns, Not significant. *$p < 0.05$. **$p < 0.01$. ****$p < 0.001$. Results significant at the Bonferroni threshold are bolded.

limitation because relying on only two measures cannot capture the full complexity of ToM. Including a wider range of tasks, such as the Strange Stories task (Happé, 1994) for cognitive ToM and the Movie for Assessment of Social Cognition (Dziobek et al., 2006) for affective ToM, would better capture different components of ToM in real-world social situations. Replicating our study with these diverse measures could also provide insights into other related ToM skills, such as hidden emotion and sarcasm, that were not observed here. That said, most tasks in the literature are not appropriate to measure ToM from young childhood to old age. We encourage future researchers to incorporate additional measures of ToM to replicate and extend our findings. We also urge researchers to develop more tasks that can measure ToM in preschoolers through older adults.

Another limitation of the present study was the use of the RMET to measure affective ToM. Psychometric research has raised concerns about the task's latent structure, failing to identify a well-fitting unidimensional or multidimensional factor structure (Higgins et al., 2023). Internal consistency, typically measured using Cronbach's alpha, has also varied widely across studies (Kittel et al., 2022). Furthermore, the RMET has limited sensitivity for discriminating among individuals with average to high levels of ToM ability; thus, it may not be an appropriate measure for non-clinical samples (Black, 2019). Relatedly, Oakley et al. (2016) showed that alexithymia (an impairment of facial recognition that co-occurs in autism spectrum disorder) accounts for differences between autism spectrum disorder and control subjects on the RMET. The authors suggest that the RMET assesses emotion recognition rather than ToM ability. Additionally, a systematic review highlighted that many studies using the RMET lack sufficient evidence of construct validity, raising concerns about the reliability of existing findings (Higgins et al., 2024). While we acknowledge these possibilities, we chose to include the RMET because it is widely used in the existing literature as a measure of affective ToM. Indeed, Baron-Cohen et al. (2001) argue that a relevant social context must be referenced from memory to understand the emotion. Further, populations with ToM deficits who score lower on the RMET compared to typically developing controls have shown comparable scores on measures of basic emotion labeling and gender-recognition control tasks (Baron-Cohen et al., 2001, 1997a,b). Ultimately, the question of whether the RMET measures ToM or emotion recognition is an important one but is beyond the scope of this work.

Finally, we acknowledge that some of our age group sample sizes were small, particularly for older adults. Thus, we had limited power to detect subtle age-related differences. Using G*Power analysis (Faul et al., 2007), we conducted sensitivity analyses for each pairwise comparison between the 65+ group and other age groups. These analyses revealed that statistical power was low across all such comparisons, indicating that our study was not sufficiently powered to detect small or even medium-sized effects involving older adults. As such, the findings related to this group should be interpreted with caution. Nonetheless, given the scarcity of longitudinal research on ToM in adults, our findings provide a valuable foundation for future studies to build from.

Conclusion

This study explored the developmental trajectory of ToM across the lifespan to explore age-related changes in ToM ability, both within and across age groups. We included separate measures of cognitive and affective ToM and used the same tasks across different age groups, spanning preschool to older adulthood. Our findings suggest that both cognitive and affective ToM remain relatively stable across the child to older adult lifespan. For both ToM components, the most pronounced developmental changes occurred during childhood, with younger children showing poorer ToM abilities compared to adults.

While this study addresses gaps in the ToM literature by using consistent measures across a diverse set of age groups, limitations, such as the reliance on a limited set of tasks and concerns about task validity, highlight the need for further research. Our results need to be replicated using more diverse methodologies. Nevertheless, our results add to a growing body of literature showing similar, yet distinct developmental trajectories for cognitive and affective ToM. Moreover, our work highlights the value of continuous ToM measures in capturing subtle changes across the lifespan.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Kwantlen Polytechnic University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

HE: Conceptualization, Writing – original draft, Writing – review & editing. RD: Formal analysis, Methodology, Writing – review & editing. DD: Conceptualization, Methodology, Supervision, Writing – review & editing. EM: Formal analysis, Methodology, Writing – review & editing. DB: Conceptualization, Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by grants from the Canada Research Chairs Program (950-232078), the Social Sciences and Humanities Research Council of Canada (435-2015-0721), and the Knowledge Mobilization Fund from Kwantlen Polytechnic University.

Conflict of interest

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1549378/full#supplementary-material

References

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., and Todd, S. (2011). Developmental continuity in theory of mind: speed and accuracy of belief-desire reasoning in children and adults. *Child Dev.* 82, 1691–1703. doi: 10.1111/j.1467-8624.2011.01635.x

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997a). Another advanced test of theory of mind: evidence from very high functioning adults with autism or asperger syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The "reading the mind in the eyes" test revised version: a study with normal adults, and adults with asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., and Jolliffe, T. (1997b). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or asperger syndrome. *Vis. Cogn.* 4, 311–331. doi: 10.1080/713756761

Bernstein, D. M. (2021). Hindsight bias and false-belief reasoning from preschool to old age. *Dev. Psychol.* 57, 1387–1402. doi: 10.1037/dev0001226

Bernstein, D. M., Coolin, A., Fischer, A. L., Thornton, W. L., and Sommerville, J. A. (2017). False-belief reasoning from 3 to 92 years of age. *PLoS ONE* 12:e0185345. doi: 10.1371/journal.pone.0185345

Bernstein, D. M., Thornton, W. L., and Sommerville, J. A. (2011). Theory of mind through the ages: older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task. *Exp. Aging Res.* 37, 481–502. doi: 10.1080/0361073X.2011.619466

Black, J. E. (2019). An IRT analysis of the reading the mind in the eyes test. *J. Pers. Assess.* 101, 425–433. doi: 10.1080/00223891.2018.1447946

Bloom, P., and German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77, B25–B31. doi: 10.1016/S0010-0277(00)00096-2

Bottiroli, S., Cavallini, E., Ceccato, I., Vecchi, T., and Lecce, S. (2016). Theory of mind in aging: comparing cognitive and affective components in the faux pas test. *Arch. Gerontol. Geriatr.* 62, 152–162. doi: 10.1016/j.archger.2015.09.009

Charlton, R. A., Barrick, T. R., Markus, H. S., and Morris, R. G. (2009). Theory of mind associations with other cognitive functions and brain imaging in normal aging. *Psychol. Aging* 24:338. doi: 10.1037/a0015225

Cho, I., and Cohen, A. S. (2019). Explaining age-related decline in theory of mind: evidence for intact competence but compromised executive function. *PLoS ONE* 14:222890. doi: 10.1371/journal.pone.0222890

Cornaggia, A., Bianco, F., Castelli, I., and Belacchi, C. (2024). Assessing metarepresentational abilities in adolescence: an exploratory study on relationships between definitional competence and theory of mind. *Front. Psychol.* 15:1456432. doi: 10.3389/fpsyg.2024.1456432

Derksen, D. G., Hunsche, M. C., Giroux, M. E., Connolly, D. A., and Bernstein, D. M. (2018). A systematic review of theory of mind's precursors and functions. *Zeitschrift Für Psychologie* 226, 87–97. doi: 10.1027/2151-2604/a000325

Dumontheil, I., Apperly, I. A., and Blakemore, S.-J. (2010). 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

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., et al. (2006). Introducing MASC: a movie for the assessment of social cognition. *J. Autism Dev. Disord.* 36, 623–636. doi: 10.1007/s10803-006-0107-0

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

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

Haskaraca, F. N., Proft, M., Liszkowski, U., and Rakoczy, H. (2023). How robust are egocentric and altercentric interference effects in social cognition? A test with explicit and implicit versions of a continuous false belief task. *Front. Psychol.* 14:1142302. doi: 10.3389/fpsyg.2023.1142302

Henry, J. D., Phillips, L. H., Ruffman, T., and Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychol. Aging* 28, 826–839. doi: 10.1037/a0030677

Higgins, W. C., Kaplan, D. M., Deschrijver, E., and Ross, R. M. (2024). Construct validity evidence reporting practices for the reading the mind in the eyes test: a systematic scoping review. *Clin. Psychol. Rev.* 108:102378. doi: 10.1016/j.cpr.2023.102378

Higgins, W. C., Ross, R. M., Langdon, R., and Polito, V. (2023). The "Reading the mind in the eyes" test shows poor psychometric properties in a large, demographically representative U.S. sample. *Assessment* 30, 1777–1789. doi: 10.1177/10731911221124342

Keysar, B., Lin, S., and Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition* 89, 25–41. doi: 10.1016/S0010-0277(03)00064-7

Kikuno, H., Mitchell, P., and Ziegler, F. (2007). How do young children process beliefs about beliefs?: evidence from response latency. *Mind Lang.* 22, 297–316. doi: 10.1111/j.1468-0017.2007.00310.x

Kittel, A. F. D., Olderbak, S., and Wilhelm, O. (2022). Sty in the Mind's eye: a meta-analytic investigation of the nomological network and internal consistency of the "Reading the mind in the eyes" test. *Assessment* 29, 872–895. doi: 10.1177/1073191121996469

Kong, Q., Selwood, O., Mfumu-Nsuka, T., Buchanan, M., Eriksen, T., Fraser, H., et al. (2025). Theory of mind development in middle childhood and adolescence: Accuracy and preference for mental-state reasoning. *J. Cogn. Dev.* 26, 274–298. doi: 10.1080/15248372.2024.2418631

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., and Moscovitch, M. (2002). Aging and autobiographical memory: dissociating episodic from semantic retrieval. *Psychol. Aging* 17, 677–689. doi: 10.1037/0882-7974.17.4.677

Lurz, R. W., Krachun, C., Mareno, M. C., and Hopkins, W. D. (2022). Do chimpanzees predict others' behavior by simulating their beliefs? *Anim. Behav. Cognit.* 9, 153–175. doi: 10.26451/abc.09.02.01.2022

Mahy, C. E. V. (2018). Mind reading across the lifespan: current themes and future challenges. *Zeitschrift Für Psychologie* 226, 134–135. doi: 10.1027/2151-2604/a000329

Meinhardt-Injac, B., Daum, M. M., and Meinhardt, G. (2020). Theory of mind development from adolescence to adulthood: testing the two-component model. *Br. J. Dev. Psychol.* 38, 289–303. doi: 10.1111/bjdp.12320

Miller, S. A. (2022). *Advanced Theory of Mind* (1st ed.). Oxford: Oxford University Press. doi: 10.1093/oso/9780197573174.001.0001

Myers, L. J., Well, D. A., and Lorch, F. R. (2010). *One-factor repeated designs* (Ed.), *Research design and statistical analysis.* (Third edition.). New York, NY: Routledge, 332–361.

Oakley, B. F. M., Brewer, R., Bird, G., and Catmur, C. (2016). Theory of mind is not theory of emotion: a cautionary note on the reading the mind in the eyes test. *Journal of Abnormal Psychology* 125, 818–823. doi: 10.1037/abn0000182

O'Connor, R. J., Lucas, A. P., and Riggs, K. J. (2024). Rethinking egocentric bias: a computer mouse-tracking study of adult belief processing. *J. Exp. Psychol. Gen.* 153, 307–327. doi: 10.1037/xge0001485

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., and Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Front. Psychol.* 6:1503. doi: 10.3389/fpsyg.2015.01503

Peterson, C. C., and Wellman, H. M. (2019). Longitudinal theory of mind (ToM) development from preschool to adolescence with and without ToM delay. *Child Dev.* 90, 1917–1934. doi: 10.1111/cdev.13064

Phillips, L. H., Bull, R., Allen, R., Insch, P., Burr, K., and Ogg, W. (2011). Lifespan aging and belief reasoning: influences of executive function and social cue decoding. *Cognition* 120, 236–247. doi: 10.1016/j.cognition.2011.05.003

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/S0140525X00076512

Raimo, S., Cropano, M., Roldán-Tapia, M. D., Ammendola, L., Malangone, D., and Santangelo, G. (2022). Cognitive and affective theory of mind across adulthood. *Brain Sci.* 12:899. doi: 10.3390/brainsci12070899

Rönnlund, M., Nyberg, L., Bäckman, L., and Nilsson, L. (2005). Stability, growth, and decline in adult life span development of declarative memory: cross-sectional and longitudinal data from a population-based study. *Psychol. Aging* 20, 3–18. doi: 10.1037/0882-7974.20.1.3

Samuel, S., Legg, E. W., Lurz, R., and Clayton, N. S. (2018). The unreliability of egocentric bias across self-other and memory-belief distinctions in the sandbox task. *R. Soc. Open Sci.* 5:181355. doi: 10.1098/rsos.181355

Shahaeian, A., Peterson, C. C., Slaughter, V., and Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Dev. Psychol.* 47, 1239–1247. doi: 10.1037/a0023899

Shamay-Tsoory, S. G., and Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia* 45, 3054–3067. doi: 10.1016/j.neuropsychologia.2007.05.021

Sommerville, J. A., Bernstein, D. M., and Meltzoff, A. N. (2013). Measuring beliefs in centimeters: private knowledge biases preschoolers' and adults' representation of others' beliefs. *Child Dev.* 84, 1846–1854. doi: 10.1111/cdev.12110

Tousignant, B., Sirois, K., Achim, A. M., Massicotte, E., and Jackson, P. L. (2017). A comprehensive assessment of social cognition from adolescence to adulthood. *Cogn. Dev.* 43, 214–223. doi: 10.1016/j.cogdev.2017.05.001

van der Wel, R. P. R. D., Sebanz, N., and Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition* 130, 128–133. doi: 10.1016/j.cognition.2013.10.004

Wellman, H. M. (2014). *Making Minds: How Theory of Mind Develops* (1st ed.). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199334919.003.0015

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

Wellman, H. M., Fang, F., Liu, D., Zhu, L., and Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychol. Sci.* 17, 1075–1081. doi: 10.1111/j.1467-9280.2006.01830.x

Wellman, H. M., Fang, F., and Peterson, C. C. (2011). Sequential progressions in a theory-of-mind scale: longitudinal perspectives. *Child Dev.* 82, 780–792. doi: 10.1111/j.1467-8624.2011.01583.x

Wellman, H. M., and Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Dev.* 75, 523–541. doi: 10.1111/j.1467-8624.2004.00691.x

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13:103. doi: 10.1016/0010-0277(83)90004-5

Yu, C., and Wellman, H. M. (2023). Where do differences in theory of mind development come from? an agent-based model of social interaction and theory of mind. *Front. Dev. Psychol.* 1:1237033. doi: 10.3389/fdpys.2023.1237033

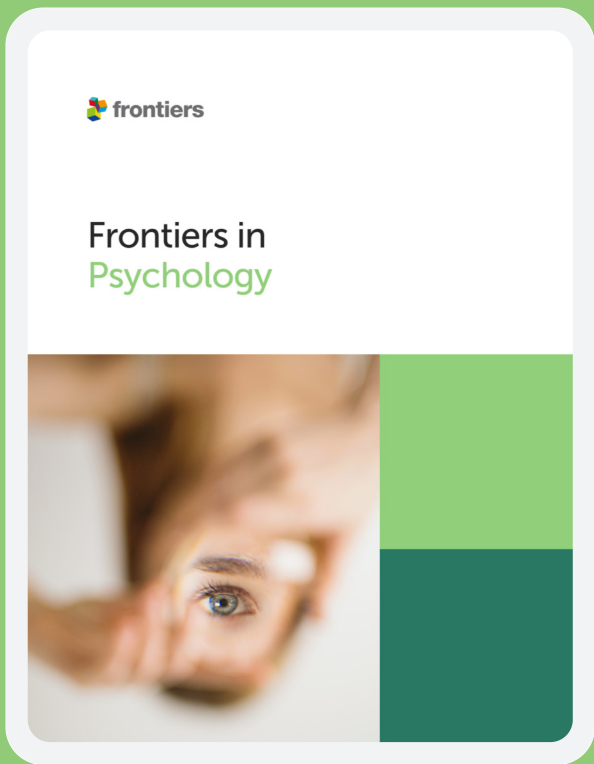
Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact