



OPEN ACCESS

EDITED BY

Shelia M. Kennison,
Oklahoma State University, United States

REVIEWED BY

Valerie Williams-Sanchez,
Valorena Online, LLC, United States
Carlo Vreden,
Leibniz Institute for Research and Information
in Education (DIPF), Germany

*CORRESPONDENCE

Debbie Kwan
✉ nyk26@cam.ac.uk

RECEIVED 03 November 2025

REVISED 05 January 2026

ACCEPTED 16 January 2026

PUBLISHED 10 February 2026

CITATION

Kwan D, Barker B, O'Farrelly C and
Ramchandani P (2026) Measuring cognitive
empathy using story stem narratives.
Front. Dev. Psychol. 4:1738605.
doi: 10.3389/fdpys.2026.1738605

COPYRIGHT

© 2026 Kwan, Barker, O'Farrelly and
Ramchandani. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Measuring cognitive empathy using story stem narratives

Debbie Kwan*, Beth Barker, Christine O'Farrelly and
Paul Ramchandani

Faculty of Education, PEDAL Research Centre, University of Cambridge, Cambridge, United Kingdom

The measurement of empathy is a challenge, particularly in young children. This study aims to develop a novel measure of cognitive empathy using story stem narratives. Children between 6 and 9 years of age ($N = 200$) completed two story stem tasks—Hurt Knee and Three's a Crowd—which were used to develop this new measure. Most children were from a white (63.5%) or mixed heritage background (20%), with near equal numbers of male (51.5%) and female (48.5%) child participants. To validate this measure, we tested for psychometric properties including reliability, internal consistency, and concurrent validity. As expected, girls scored higher for cognitive empathy than boys (mean difference = 1.15, 95% CI [0.46, 1.84]). Associations were seen between cognitive empathy and parent-reported prosocial behavior ($r = 0.16$, $p = 0.02$), and callous-unemotional traits ($r = -0.15$, $p = 0.02$). Using children's representations of their world through story stem narratives could allow for new insights in the study of empathy in young children.

KEYWORDS

cognitive empathy, empathy, measure of empathy, social and emotional development, story stem narratives

Introduction

Empathy is an important component in our ability to socialize and sustain relationships with others. Research has consistently shown positive associations between empathy and prosocial behavior, particularly in studies of helping. Prosocial behavior is defined as voluntary behavior that is intended to benefit another (Eisenberg et al., 2006). While helping falls under prosocial behavior, it is more specifically defined by actions that provide assistance and/or resources to benefit a needy other (Eisenberg and Mussen, 1989). In studies involving adults and children, there is substantial evidence to suggest that empathic individuals are likely to be more helpful [see for example, a recent meta-analysis by Qiu et al. (2024)].

Beyond benefits to others in the form of prosociality, empathy is also beneficial to the self, as it has been linked to social competence (Saarni, 1990; Eisenberg and Miller, 1987), peer likeability (Denham et al., 1990), a reduced risk in loneliness and depression (Beadle et al., 2012; Tully et al., 2016), and even in the promotion of creativity (Demetriou and Nicholl, 2022).

While empathy's positive role in social and emotional development is hardly disputed, there remain many challenges in the measurement of empathy, particularly in young children (Wang and Wang, 2015). Additionally, disparities and inconsistencies in the definition of empathy and its closely related constructs, like sympathy, have led researchers to employ a wide range of measures of empathic behavior (Pederson, 2009; Gerdes et al., 2010).

The construct of empathy

While extensively studied, the concept of empathy has been difficult to define. Researchers have previously fallen into two groups, where some have viewed empathy as an affective state, while others have taken a more cognitive approach to its interpretation (Baron-Cohen and Wheelwright, 2004). More specifically, affective empathy refers to the sharing of emotions and has been synonymously linked to emotion contagion, as one experiences the same emotions or physiological states as somebody else (Hsee, 1990; Decety and Holvoet, 2021). In contrast, cognitive empathy denotes the understanding of what another person is feeling (Trimmer et al., 2017).

There is now growing consensus in current literature that empathy is made up of both a cognitive and an affective component (Simon and Nader-Grosbois, 2023). A fully empathic response will depend not only on the ability to understand another's emotional state, but it will also involve the observer's own emotional reaction to the stimulus (Davis, 1980; Eisenberg et al., 2006; Batson, 2009; Roth-Hanania et al., 2011).

Measures of empathy

Self-reported measures

Studies of empathy have typically used self-reported measures to assess empathy in individuals, the most popular being Mehrabian and Epstein's (1972) Questionnaire Measure of Emotional Empathy (QMEE), Bryant's Index of Empathy (Bryant, 1982), and the Interpersonal Reactivity Index (IRI; Davis, 1983). The QMEE was developed to measure emotional (affective) empathy in adults, with 33 questions falling under the subscales of "Susceptibility to Emotional Contagion", "Appreciation of the Feelings of Unfamiliar and Distant Others", "Extreme Emotional Responsiveness", and "Tendency To Be Moved by Others' Positive Emotional Experiences". Responses to each item ranges from +4, suggesting very strong agreement, to -4, indicating very strong disagreement. Bryant's Index of Empathy (Bryant, 1982) adapts the QMEE (Mehrabian and Epstein, 1972) to be suitable for use with school-aged children and adolescents. This new scale consists of 22 questions where respondents are asked to answer questions with a yes/no format. While the QMEE and Bryant's Index of Empathy primarily focus on the affective component of empathy, the IRI (Davis, 1983) encompasses both affective and cognitive components. The IRI consists of 28 items that fall under four different subscales, that is, "Perspective-Taking", "Fantasy", "Empathic Concern", and "Personal Distress". Respondents are required to rate the questions on a 5-point Likert scale ranging from "Does not describe me well" to "Describes me very well".

Conceptual limitations have been evidenced in the QMEE and IRI, as questions measuring affective empathy are perceived reactions, rather than shared emotions (Jolliffe and Farrington, 2006). Most self-reported measures are also created for an adolescent and/or adult population and are not suitable for young children. Although the Index of Empathy (Bryant, 1982) provides a scale for a younger population (6+ years), a systematic review (Sesso et al., 2021) showed this measure to have one of the lowest

internal consistencies out of the 16 measures in review. Further to the Index of Empathy (Bryant, 1982), Bensalah et al. (2016) presents a more recent self-reported measure suitable for middle-aged children (6–12 years old). This measure is adapted from the adult version of the Basic Empathy Scale (BES; Jolliffe and Farrington, 2006) and contains 20 items involving cognitive and affective empathy that are rated on a 5-point Likert scale. While similarities between the adult version (BES-A) and the child version (BES-C) are acknowledged, highlighting the scale's potential for studying empathy in middle childhood, the authors also indicate a limitation in the scale's low internal consistency.

Beyond the issues discussed, there are other validity concerns worth acknowledging in self-reported measures (Eisenberg and Miller, 1987). Specifically, reported answers can be affected by social desirability (Neumann et al., 2015), language ability (Wang and Wang, 2015), and even more fundamentally, whether young children are aware of what they are feeling and can/are willing to report their emotional states accurately (Eisenberg and Fabes, 1990). Thus, further investigations are needed.

Picture-story methods have also been used to measure empathy in children, with the most widely used being Feshbach and Roe's Affective Situations Test for Empathy (FASTE; Feshbach and Roe, 1968; Zhou et al., 2003). This measure involves the child being told a story while shown pictures of protagonists in emotionally charged situations (Zhou et al., 2003) and then being asked how the protagonist feels and how they, themselves, feel (Feshbach and Roe, 1968). Although the FASTE was an important early method (Eisenberg-Berg and Lennon, 1980; Iannotti, 1985), this measure captures only the affective component of empathy. Furthermore, concerns lie in the stories ability to evoke sufficient affective responses (Lennon et al., 1983; Eisenberg and Miller, 1987), alongside social demands (Eisenberg-Berg and Lennon, 1980) and gender-biases (Eisenberg and Lennon, 1983; Lennon et al., 1983).

A more recent performance-based measure called the Kids' Empathic Development Scale (KEDS; Reid et al., 2013) developed for 7–11-year-olds, also exposes children to emotional stories but is inclusive of a cognitive component. Children are asked to provide a justification that explains their emotional inferences. However, this measure has elicited its own criticisms. Particularly, picture scenarios may restrict the depth and capacity for which perspective-taking is measured, as cognitive empathy should relate to the understanding of emotional situations that are specific to people's emotional states (Bensalah et al., 2016). The affective component of the measure has also been criticized, as it requires children to draw on an emotional inference, rather than the explicit sharing of emotions (Bensalah et al., 2016). Overall, self-reported measures whether conducted through questionnaires or picture-story methods each have their existing limitations that need to be considered further.

Observational measures

Researchers have developed observational methods for measuring empathy, which assess children's facial, gestural, and behavioral responses (Eisenberg and Fabes, 1990). Generally, following Zahn-Waxler et al. (1992) observational coding scheme, infants' reactions to their mother's and/or an experimenter's

distress is measured. Due to its ability to reduce self-presentational biases and its capacity to lessen language comprehension barriers (Zhou et al., 2003), studies have been able to measure empathy through observational approaches in children and in infants as young as 15 months old (see Miller et al., 1996; Zahn-Waxler et al., 1992). In more recent studies, rare cases of comforting, as well as empathic concern have also been evidenced in infants as young as 8 or 9 months old (see Roth-Hanania et al., 2011; Vredén et al., 2025).

While observational measures may provide a more objective measure of children's empathy, they still present limitations. Arguably, they do not fully capture the cognitive component of empathy. Although children's exploration or attempt to understand a distressed victim, known as 'hypothesis testing,' or 'inquiry behavior,' gives a signal that they are trying to understand the situation, this behavior does not indicate what the child actually understands (Davidov et al., 2021). Thus, hypothesis testing, as captured under the cognitive component of empathy in existing observational measures (Zahn-Waxler et al., 1992) should not be used as a pure assessment of cognitive empathy (Davidov et al., 2021).

In sum, the measurement of empathy remains a challenge. Administration, language barriers, validity of scales, and the lack of cognitive and affective aspects have each presented a limitation to the methodological approaches used to measure empathy. Self-reported measures, though frequently used in studies of adults and children, can present biases that affect the accuracy of participants' answers. While observational measures seem to reduce biases, the cognitive component of empathy cannot be inferred from behavior alone (Hoffman, 2000).

Story stem narratives

Narrative-based dolls play assessments encourage children to manipulate dolls and figurines while telling stories (Buchsbau and Emde, 1990). These play- and story-based methods offer a developmentally engaging way to understand the perspectives of young children. The MacArthur Story Stem Battery (Bretherton and Oppenheim, 2003; Bretherton et al., 1990) was one of the earliest systematic story stem assessment tools. In this story stems task, an adult interviewer delivers an emotionally charged story beginning, while the child participant is invited to complete the story through role-play and manipulation of the dolls (Yuval-Adler and Oppenheim, 2014). Although largely representational, this technique offers a unique and unrestricted perspective of children's behaviors (Arseneault et al., 2005; Kraemer et al., 2003) in response to scenarios that children themselves will have typically, experienced before. The verbal skills required for narrative representations are generally evident by 3 years of age, suggesting that such methodologies are appropriate for use in young children (Bettmann and Lundahl, 2007; Buchsbau and Emde, 1990).

This storytelling approach to understand children's experiences and representations has its roots in clinical and attachment-based research (Bretherton and Oppenheim, 2003). Bowlby (1940, 1944, 1958) and Ainsworth (1970) demonstrated that young children form mental schemas about interpersonal relationships according

to past experiences, and that their social and emotional functioning is closely linked to these schemas (Kobak, 1999). Story stems were, therefore, a method primarily used to understand children's internal worlds within mental health research. Although story stem tasks often explore themes rich in emotional content (e.g., mother-infant interactions, social competence, moral values, family conflict, etc.), it is not commonly used to measure empathy, representing a missed opportunity in developmental research. The few papers that measure empathy using story stems (see Petrowski et al., 2009, 2014) follow the MacArthur Narrative Coding System (MNCS; Robinson et al., 1992). While the MNCS provides a strong foundation, the coding of empathy within this scheme can be quite broad. For example, the same behavior that is coded within the Empathic Relations theme (e.g., sharing, helping, affiliation, affection, and/or reparation/guilt), can also be tracked under other themes, like Parental Positive Representations, Interpersonal Conflict, and Avoidance/Dissociation.

Story stem narratives are a valuable means of exploring empathy in young children. Firstly, the playful nature of this storytelling approach makes this measurement developmentally appropriate and authentic. Secondly, the stories presented are situations which children will have typically experienced before, allowing children to demonstrate how they would respond to familiar situations, given previous experiences. Finally, story stem tasks present emotionally charged situations, which are crucial to capturing empathy, as one must be able to understand the feelings of other people. Taken together, story stem narratives may overcome some of the limitations reported earlier in existing observational and self-reported methods. Despite story stems being a potentially instructive way to measure empathy in young children, there does not appear to be any specific story stem task that can/is used exclusively for measuring empathy. The breadth of this work can make it difficult for researchers to use this measurement approach in a standardized and reliable manner. Thus, the development of a simpler coding approach, using specified story stem tasks to measure empathy could improve consistency in the way we can measure empathy using story stem narratives.

The present study

The aim of this study is to develop a measure of cognitive empathy, focusing on two specific story stem scenarios—Hurt Knee (Dolls House Play: Woolgar and Murray, 2010) and Three's a Crowd (MacArthur Story Stem Battery: Bretherton and Oppenheim, 2003). By developing this coding scheme using two specific story stem narratives, this work aims to contribute to the literature by putting forward a child-appropriate, yet easy to administer task that can capture, specifically, children's cognitive empathy. This play-based measurement tool is intended to supplement other observational approaches of affective empathy to capture both components.

Research aims

The primary aim of this exploratory study is to develop and test the psychometric properties of a new coding scheme developed

TABLE 1 Descriptive statistics for study variables.

Variable	N	M	SD	Median	Range	IQR	Skewness	Kurtosis
Story stem measures								
Cognitive empathy	200	6.73	2.53	7.00	12.00	2.00	-0.48	3.04
Prosocial behavior	185	5.16	1.83	5.00	9.00	3.00	-0.45	2.62
Parent-reported measures								
SDQ—Prosocial Behavior	200	7.96	1.90	8.00	10.00	3.00	-0.91	3.44
Callous-unemotional traits scale	195	3.20	2.24	3.00	11.00	3.00	0.73	3.21
Teacher-reported measures								
SDQ—Prosocial Behavior	97	7.71	2.32	8.00	10.00	4.00	-0.84	2.81

SDQ, Strengths and Difficulties Questionnaires.

to measure cognitive empathy using story stem narratives. Specifically, to:

1. Explore inter-rater reliability and internal consistency.
2. Explore the concurrent associations between empathy and closely related constructs such as prosocial behavior and callous-unemotional traits, as well as child gender and age, when children are 6–9 years of age.

Although this work is exploratory, predictions are made for the following associations. Based on previous studies, it was hypothesized that cognitive empathy and prosocial behavior will show positive associations (Spinrad and Eisenberg, 2014; Qiu et al., 2024). As parent and teacher reports are used to measure children's prosocial behavior, we also suspect some differences may surface, given that the frequency (Qiu et al., 2024) in which children demonstrate prosocial behaviors can differ by context (Deneault et al., 2023), and to whom behaviors are being displayed (e.g., friends, siblings, adults, strangers, etc.). We also hypothesized that cognitive empathy and CU traits would show a negative association, in line with previous research (Lui et al., 2016; Georgiou et al., 2019; Waller et al., 2020). Furthermore, we hypothesized a positive association between cognitive empathy and gender, with girls typically demonstrating higher empathic behavior than boys (Petrowski et al., 2014; Basay et al., 2021; Sultan and Khan, 2025). Finally, we suspect there may be no association between cognitive empathy and age in the current study, as our age range is relatively small (i.e., 6–9 years). Previous studies during key developmental periods have yielded significant differences in empathy as a function of age (Hoffman, 1990), for example, changes in empathy from infancy through to middle childhood. However, studies exploring smaller age ranges have also evidenced non-significant results (Zajdel et al., 2013).

Method

Study design and participants

This study uses secondary data from the Healthy Start, Happy Start project (see O'Farrelly et al., 2021 for full recruitment process, participant information, and study procedure). A total of 300 children and their primary caregivers were recruited predominantly from health visiting services based on children

scoring in the top 20% on the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). Research assessments were conducted in participants' homes (or, in a very small number of cases, in private rooms of children's centers). Trained researchers collected family demographic information and conducted the tests, typically, within 2 h during these home visits. While primary caregivers were completing the measures, the story stem tasks were conducted with the child participant. These interactions were video recorded and later coded. Children were aged between 1 and 3 years and were followed up for 6 years. This study uses data at the 6-year follow up, when children are between 6 and 9 years of age. This age group was chosen, as most children can pass basic cognitive empathy/theory of mind tests by 5–6 years of age (O'Reilly and Peterson, 2015), with evidence to suggest that cognitive abilities continue to develop throughout middle childhood (Devine et al., 2016; Devine and Hughes, 2016). Informed consent was provided by the parent/caregiver at each stage of the study, while participating children were also invited to take part in a child assent procedure. Ethical approval was granted by Riverside Research Ethics Committee (REC ref: 14/LO/2071). The trial was funded by the National Institute for Health Research, Health Technology Assessment (NIHR HTA; ref: 13/04/33). See Table 1 for descriptive statistics of the study variables.

Primary outcomes

Child informant measure: story stem narratives

Two story stems, Hurt Knee (Woolgar and Murray, 2010) and Three's a Crowd (Bretherton and Oppenheim, 2003) were chosen, as each scenario consists of an emotionally charged story beginning that involves a child being put in a position of potential distress or needing care. These scenarios present the child participant with an opportunity to elicit a possible empathic response toward the protagonist in the story. Children are invited to choose dolls that represent themselves and their family members, as a means of accessing their representations of their relationships and their own and others' behaviors.

Hurt Knee Description. In these examples the participant child is called Sarah, and her friend is called Molly. The experimenter says, "it's a new day and the family are at the park. Mummy and daddy are sitting over here. Sarah is playing with Molly. Sarah

TABLE 2 Scale construction and subscales with examples for hurt knee.

Subscales	Examples
Care/concern	
C1 = Awareness of hurt child.	“Are you okay?” “What happened?”
C2 = Active concern or care.	“You shouldn’t run again.” “I think you should have a rest.”
Helping	
H1 = Seeks help from parent/adult but must be relevant to resolving the situation (e.g., if mum says, “it’s time to go home now”, this behavior is not counted).	“Mum wipes her knee.” “Dad puts a plaster on.”
H2 = Helping child directly through physical behavior.	“They all go to the hospital.” “X helps Y get back up.”
Perspective-taking	
P1 = Following prompts by experimenter (“how does X feel, and why does X feel this way?”).	“X feels sad. She feels sad because she cut her knee.” “Mum feels worried. She feels worried because her child is hurt.”
P2 = Spontaneous perspective-taking during the doll play.	“X feels better so he keeps playing.” “She starts crying because it still really hurts.”

runs, falls, and cuts her knee. Sarah says, “ow! I’ve really hurt my knee!” Show me and tell me what happens next”. While the child is manipulating the dolls to finish the story, the experimenter will also ask, “how does Sarah feel when she fell over? Why does she feel this way?” and “how does mummy feel when Sarah fell over? Why does she feel this way?”. These questions are only asked of the child participant doll who falls over, and of the primary caregiver doll who is witnessing the situation.

Three’s a Crowd Description. In these examples, the participant child is called Sarah. Ava is her friend, while Sophie is another child that is brought into the story. The experimenter says, “it’s a new day and the family are at another park. Mummy and daddy are sitting over here. Sarah is playing with Ava and her new ball. Sophie comes along and asks if she can play. Sarah says, “sure”, but Ava says, “no way, I don’t want to play with Sophie”. Show me and tell me what happens next”. While the child is manipulating the dolls to finish the story, the experimenter will also ask, “how does Sarah feel when Sophie couldn’t play? Why does she feel this way?” and “how does Sophie feel when she couldn’t play? Why does she feel this way?”. These questions are only asked of the child participant doll who is witnessing the situation, and of the child doll who could not play.

Development of coding scheme

Coding of cognitive empathy

Coding of cognitive empathy is captured by three categories in each of the two tasks, that is, whether the doll play shows (1) helping, (2) care/concern, and (3) perspective-taking. Within each

TABLE 3 Scale construction and subscales with examples for three’s a crowd.

Subscales	Examples of Participant Responses
Care/concern	
C1 = Awareness of excluded child.	“Hey, that’s not nice.” “Why don’t you want to play with them?”
C2 = Active concern or care – sticking up for excluded child.	“I won’t play if you don’t let him play.” “Come on, we need one more player anyway.”
Helping	
H1 = Seeks help from parent/adult but must be relevant to resolving the situation (e.g., if mum says, “it’s time to go home now”, this behavior is not counted).	“Mum comes over and tells him off.” “Dad says she should let them play.”
H2 = Helping child be included through physical behavior.	“Now all three of them play together.” “Fine, I’ll play with you then.”
Perspective-taking	
P1 = Following prompts by experimenter (“how does X feel, and why does X feel this way?”).	“X feels sad. He feels sad because the other boy wouldn’t let him play and it wasn’t kind.” “Y feels disappointed. He feels disappointed because he really wanted to play.”
P2 = Spontaneous perspective-taking during the doll play.	“X feels happy because now they’re all playing together.” “Y says sorry for being mean, and then X feels good again.”

of the three categories, there are two subscales of behaviors that can exist. See Tables 2, 3 for the subscales and examples of participant responses in each of the two tasks.

These subscales were created with the intention of capturing a variation of behaviors possible, as shown by the child participants. Under helping, the child could seek help from an adult or show direct helping through physical behavior in the doll play. Under care/concern, the child may verbalize comments that show care or concern (e.g., “are you okay”, “just sit down now and rest”). Under perspective-taking, the participant child must correctly answer the experimenter’s questions of how particular characters are feeling and why, or they must state character’s emotions during the doll play. These behaviors are coded as present (1) or absent (0). A final empathy score is given at the end of both tasks, by summing all the scores of the categories together. Ultimately, the maximum score a child can receive is 12, representing the highest level of cognitive empathy in this coding scheme. See Appendix A for an example of the coding, and the Supplementary materials for a detailed description.

The three categories

The three categories of helping, care/concern and perspective-taking were generated through exploring existing measures combined with literature, and viewing the data repeatedly, to see

what responses children were showing during each of the story stem tasks.

Observational studies of empathy in existing literature have commonly used Zahn-Waxler et al. (1992) version of the MacArthur Longitudinal Twin Study coding scheme (Emde et al., 1992). Many of the responses coded involved facial, behavioral, or gestural expressions for the victim (e.g., sustained expressions, brows furrowing, crying, etc.). However, the behaviors shown in the story stem tasks did not elicit such affective responses. Cognitive empathy is, therefore, the focus of this measure.

Helping. Helping is defined by actions that provide assistance and/or resources to benefit a needy other (Eisenberg and Mussen, 1989). This category draws influence from the MNCS, as prosocial behaviors like helping, are categorized under the Empathic Relations theme. In the Hurt Knee task, children often demonstrated helping behaviors, for example, when a plaster was given to the hurt child, or in the Three's a Crowd task, one might be helping the excluded child be included in the play.

Care/concern. Caring is defined as having an emotional commitment to an individual's well-being that can encourage prosocial efforts to support them (Weiner and Auster, 2008). Concern for others is the tendency to express interest in others' emotional distress (Rhee et al., 2013). These behaviors were coded together, as in the Hurt Knee and the Three's a Crowd tasks, children would often express care or concern with verbal comments such as "are you okay?" or "it'll be okay, let's not run again". Drawing influence from the MNCS, behaviors demonstrating reassurance (e.g., "don't worry, I'm here") were also categorized under the Empathic Relations theme.

According to literature, prosocial behaviors should be considered as a part of the 'empathic repertoire', as children may demonstrate an emotional reaction (i.e., an attempt to alleviate the distress) that conveys some sort of understanding or concern for a victim in distress (Moreno et al., 2008).

Perspective-taking. Perspective-taking is arguably, the essential factor of cognitive empathy, as it represents one's ability to recognize and understand the feelings of others. One of the most crucial features of a story stem task is that the experimenter explicitly asks the child, "how does X feel?" and "why does X feel this way?". These questions serve to elicit cognitive understanding and allow researchers to determine whether a child recognizes how particular characters in the stories are feeling and why.

Secondary outcomes

Measure of prosocial behavior

Strengths and Difficulties Questionnaire. The Strengths and Difficulties Questionnaire (SDQ, school-aged version for 4–17 years; Goodman, 1997) has been widely used in clinical and educational research (Goodman and Scott, 1999) to understand children's behavior. This study uses only the prosocial behavior subscale of the SDQ, which contains five questions. Respondents are asked to rate each item on a three-point scale (0 = not true, 1 = somewhat true, or 2 = certainly true) according to the child's behavior within the last 6 months. Where parental consent for contact with the child's school was given, the child's

class teacher was also invited to complete the SDQ. Parent and teacher reports are widely used in studies of children's mental health and wellbeing (Murray et al., 2021), as each provides a different perspective and teachers and parents are reporting on what they see across different settings (i.e., home vs. school). Studies have reported satisfactory psychometric properties related to reliability and validity when using the SDQ (Mieloo et al., 2012).

Story Stem Narratives. Children's prosocial behavior was also coded using the story stem method. This was a novel scale developed from the initial study trial, using ideas from the MNCS but rated on a 0–3 scale to assess the intensity of behaviors shown. A different coding scheme was used but involved the same two tasks of Hurt Knee and Three's a Crowd. An additional story stem task called Spilt Juice (see The MacArthur Story Stem Battery; Bretherton et al., 1990; Bretherton and Oppenheim, 2003), was also used in this coding scheme to explore prosocial behavior. However, there was little variance in children's expression of empathy during piloting of coding using this story stem and thus, it was excluded. Prosocial behaviors under this measure were coded on a scale consisting of 0 = no prosocial behaviors, 1 = minimal, 2 = moderate, or 3 = extreme. Minimal prosocial behaviors were evidenced if characters were friendly toward each other but not going out of their way to include or help (e.g., "I clean up the juice"). Moderate behaviors were evidenced if characters were explicitly helpful, considerate, comforting or friendly (e.g., "mum helps me clean up the juice"). Extreme behaviors were evidenced if characters go above and beyond to help, or if characters modify their behavior to help (e.g., "since dad drank all the juice, I go to the shop to get him some more").

Measure of callous-unemotional traits

Callous-unemotional traits scale

Callous-unemotional traits (CU traits) were assessed using a seven-item scale, as reported by the primary caregiver. Four items were taken from the SDQ prosocial scale [i.e., 'considerate of other people's feelings' (reverse-scored)], while three items were taken from the Inventory of Callous-Unemotional Traits (ICU; Frick, 2004) [i.e., 'does not show feelings or emotions']. This combined scale has been used by previous groups and satisfactory psychometric properties have been reported (see Dadds et al., 2014; Takahashi et al., 2021).

Analytic plan

The first set of analyses explores inter-rater reliability using Cohen's Kappa (Cohen, 1960) and internal consistency using Cronbach's alpha (Cronbach, 1951). The second set of analyses uses a *t*-test to explore the association between gender and cognitive empathy, a linear regression to examine gender as a potential moderator of cognitive empathy and age, and then several bivariate (Pearson) correlations are conducted to explore the concurrent associations between common indicators associated with empathy such as prosocial behavior and callous-unemotional traits.

TABLE 4 Demographic information for participating children and their primary caregivers.

Characteristic (n%), unless stated	Full sample (N = 300)	Six-year follow-up (N = 200)
Child Characteristic		
Gender		
Female	137 (46)	97 (48.5)
Male	163 (54)	103 (51.5)
Age (years), mean (SD)		
Baseline	1.9(6.7)	-
6-year follow-up	-	8.2(0.7)
Ethnicity		
White	194(65)	127(63.5)
Mixed	61(20)	40(20)
Asian	17(6)	12(6)
Black	18(6)	13(6.5)
Other	10(3)	8(4)
Primary caregiver characteristic		
Gender		
Female	287(96)	194(97)
Male	13(4)	6(3)
Age (years), mean (SD)		
Baseline	34.2(5.8)	-
6-year follow-up	-	41.4(5.2)
Relationship to child		
Biological or adoptive mother	288(96)	194(97)
Biological or adoptive father	12(4)	6(3)
Ethnicity		
White	217(73)	140(70)
Mixed	22(7)	14(7)
Asian	31(10)	20(10)
Black	18(6)	15(7.5)
Other	12(4)	11(5.5)
Relationship status		
Married/civil partnership/cohabiting	255(85)	165(82.5)
Divorced/widowed/legally separated	5(2)	8(4)
Single and none of the above	29(10)	25(12.5)
In a relationship but not cohabiting	11(3)	2(1)
Highest qualification		
GCSE or lower	31(10)	16(8.2)
College (e.g., A levels, NVQ, BTEC)	78(26)	42(21.4)
Graduate	191(64)	138(70.4)

Results

Sample characteristics

Three hundred participants were part of the original Healthy Start, Happy Start (HSHS) trial. At the 6-year follow up, 241 children remained in the study. However, only 200 participants (an approximately 67% retention rate) completed the face-to-face story stem assessment. Full parent and child data is available for 185 participants. The teacher sample was much smaller ($n = 97$). Children in the current study were between 6.7 and 9.6 years of age ($M = 8.18$ years, $SD = 0.65$ years). There were near equal numbers of male (51.5%) and female (48.5%) child participants. Most children were from a white (63.5%) or mixed heritage (20%) background. Participating caregivers were predominantly female (97%) and were likely to be the child's biological mother (97%). In total, 70.4% of children's primary caregivers held a graduate level qualification. See Table 4 for full demographic details.

Exploring inter-rater reliability and internal consistency

Inter-rater reliability

Cohen's Kappa Cohen (1960) was used to determine the inter-rater agreement between two postgraduate researchers trained on this coding scale (double-coding 7.5% of the videos). The coders reached 0.91 agreement for the total cognitive empathy score—the main outcome of interest. Scores between 0.81 and 1.00 represent an excellent level of agreement (McHugh, 2012).

Internal consistency

Cronbach's alpha Cronbach's (1951) was used to test for internal consistency. The total cognitive empathy score had an alpha of 0.68, which reached 'acceptable' reliability, that is, within 0.6–0.7 (using the rule-of-thumb guidelines; Nunnally, 1967, 1978; Nunnally and Bernstein, 1994).

Exploring concurrent associations

Gender and age

Females scored higher than males for the overall, total cognitive empathy score (mean difference = 1.15, 95% CI [0.46, 1.84]).

There was no evidence of an association in cognitive empathy by age ($r = 0.07$, $p = 0.32$), although the range of age was not very large (6.7–9.6 years).

A moderation analysis was performed by adding an interaction term (age by gender) with cognitive empathy. This additional test was conducted to determine whether the relationship between cognitive empathy and age differs by gender. The interaction between age and gender was not statistically significant ($b = -0.69$, $p = 0.20$). This suggests that gender does not change how age relates to empathy.

Prosocial behavior

There was a small positive association found between parent-reported prosocial behavior (SDQ) and cognitive empathy ($r = 0.16$, $p = 0.02$). However, no association was found between teacher-reported prosocial behavior and cognitive empathy ($r = 0.09$, $p = 0.33$), though fewer teachers ($n = 97$) reported. A strong positive association was found between child-reported prosocial behavior and cognitive empathy using story stems ($r = 0.61$, $p = 0.00$).

Callous-unemotional traits

There was an association between callous-unemotional traits (CU traits) and cognitive empathy with a small negative correlation ($r = -0.15$, $p = 0.02$).

Discussion

In this study, we developed and tested the psychometric properties of a novel coding measure of cognitive empathy in a story stems task. This coding scheme extends previous literature by using story stem narratives to explore children's understanding of others' emotions through children's representations of the world. For purpose of validation, we analyzed the associations between our measure of cognitive empathy and various related constructs. Results demonstrated an expected difference by gender and found associations between children's cognitive empathy and prosocial behavior, and callous-unemotional traits (CU traits).

Gender and age

Gender differences were found in children's total cognitive empathy score. Aligning with existing literature, females tend to show higher levels of empathy whether in studies of children or adolescents (see for example Jolliffe and Farrington, 2011; Belacchi and Farina, 2012). In Baron-Cohen's (2002) article, studies examining sharing, rough play, empathic responding, theory of mind, sensitivity to facial expressions, values in relationships, and aggression, are summarized, each revealing sex differences of a small but significant magnitude. Specifically, girls have shown stronger concerns for fairness, earlier use of theory of mind, and a greater likelihood of endorsing cooperative norms (see Baron-Cohen, 2002). In contrast, boys tend to engage in more rough-and-tumble play, exhibit more direct aggressive behaviors (e.g., pushing, hitting), and have higher prevalence rates of conduct disorders (see Baron-Cohen, 2002). These behavioral differences may help to explain the gender differences often observed between females and males in studies of empathy.

Gender differences observed during middle childhood may also be partly explained by socialization (Brody, 2013). As children begin to internalize social roles and gender norms by referring to adult expectations (Simon and Nader-Grosbois, 2023), girls tend to show more concern for the emotions of others and act prosocially, compared to boys (Strauss, 2004; Bensalah et al., 2016). In line with the cognitive component of empathy, females between 5 and 9 years of age have shown to possess higher mentalizing abilities than males

(Hughes et al., 2005). This may account for why females in the present study were observed having higher cognitive empathy than males, as their ability to infer the emotions of others may have been more developed than that of males at this age.

The current analyses revealed no overall effect of age on cognitive empathy, nor did gender moderate this association. In existing literature, age and empathy have shown mixed results. Studies exploring a small age range have evidenced non-significant results, while significant differences as a function of age have been evidenced during key developmental periods. For example, changes in empathy between infancy and middle childhood can be expected, as perspective-taking begins to develop around 2–3 years of age (Hoffman, 1990). Likewise, between middle childhood and late childhood, changes in children's empathy can also be driven by advancing language and cognitive abilities, enabling comprehension of more complex emotions. Zajdel et al.'s (2013) study exploring children at 5–7, 8–9, and 10–12 years of age, found emotion understanding improving with age, however, only between the youngest and the oldest age group (i.e., 5–12 years). Follow-up comparisons between the 5–7- and 8–9-year-olds showed no significant differences, nor did the 8–9- and 10–12-year-old groups, respectively. In line with the age of children in the present study—6.7 and 9.6 years of age—the narrow age range interpretation may help to explain our similar result of non-association between age and cognitive empathy.

Prosocial behavior

Prosocial behavior showed a small positive association with children's cognitive empathy score when measured by the parent-reported SDQ and a strong positive association when child-reported via story stems. These findings are largely in line with previous literature, where prosocial behavior and empathy are consistently positively linked. In the case of the present study, cognitive empathy, that is, the ability to understand other's emotions, may enhance prosocial responses when children recognize when others are in distress, are needing help, or are being excluded or left out (i.e., in the context of the Hurt Knee and Three's a Crowd scenarios). Theoretically, it is compelling to assume a reciprocal link between perspective-taking abilities and prosocial behavior, as one must be able to consider the goals, needs and intentions of others (Dunfield, 2014; Hay and Cook, 2007) in order to engage in any form of prosocial, voluntary behavior. Our findings suggest that the ability to comprehend other's feelings and emotions can motivate observable prosocial behaviors even in everyday contexts. Although cognitive empathy refers to inferring the mental and emotional states of others (Trimmer et al., 2017), it can also have the capacity to elicit some physiological arousal that can motivate one to respond prosocially (Wang and Wang, 2015). Indeed, individuals who experience other-orientated emotional reactions or experience concern toward others may be more inclined to respond with prosocial acts (Spinrad and Eisenberg, 2014). While empathy does not always guarantee prosocial behavior, for example, altruistic behavior can be driven by moral values rather than empathic responding (Eisenberg-Berg, 1979; Schwartz and Howard, 1984), empirical

evidence certainly advocates a strong relationship (Soliman et al., 2021).

In contrast, we found no association with teacher-reported SDQ. We suspect that this result may be partly impacted by the response rates, as fewer teachers reported compared to parents. Another possible explanation is related to differences in the context (Deneault et al., 2023) and frequency (Qiu et al., 2024) in which prosocial behaviors are being observed by parents in a home environment vs. teachers in a school/classroom setting. It may be possible that parents have more opportunities to observe children's behaviors, or that children at risk of behavior problems may find it more challenging to act prosocially in larger group settings where social demands are higher than smaller group interactions that parents typically observe. Indeed, teachers and parents interact with the child in different ways and with different observed behaviors within different contexts. Teachers certainly have an important perspective on children's behaviors; however, it could be argued that parents provide a more holistic view of the child that spans across different environmental settings.

Callous-unemotional traits

CU traits showed a small negative association with children's empathy score in this study. This suggests that as empathy increases, callous-unemotional traits tend to decrease, and vice versa. Our result corroborates existing studies showing empathy being negatively correlated with CU traits (Waller et al., 2020). Given that CU traits are generally defined by having low empathy, a lack of remorse, and an insensitivity to the distress of others (Frick et al., 1994, 2014), it is not surprising that these behaviors are reduced in empathetic individuals who seem to perceive and resonate with the emotions felt by others (Eisenberg et al., 2015).

Limitations

Our study has some methodological limitations that warrant further exploration. The main limitation is that our study did not have another measure of cognitive empathy to compare the new measure with. This would have been useful to further validate our measure. As the current study utilizes secondary data, we were limited by the predetermined nature of the variables and measures available for analysis. Future work could expand on the current coding scheme by including other story stem tasks [e.g., "Mom's headache" (see The MacArthur Story Stem Battery; Bretherton et al., 1990; Bretherton and Oppenheim, 2003)] that would facilitate other emotional responses from child participants. Other constructs that are related to empathy could also be explored by including additional associations (e.g., executive functions, personality traits, etc.) to support the validity of this work. Another limitation is our inability to capture affective empathy in this coding scheme. Children in the current study did not seem to express any major behavioral responses in either of the story stem tasks. Thus, researchers may need to use other measures, such as experimental/observational methods (see for examples Zahn-Waxler et al., 1992; Knafo et al., 2008; Roth-Hanania et al., 2011) that capture behaviors indicative of affective empathy to

compliment the current measure. Although, affective empathy was not considered in this study, the singular exploration of cognitive empathy also has its strengths. Previous work has often neglected to make distinctions between cognitive and affective empathy, even though both factors have distinct neural and behavioral correlates (Lilienfeld, 2003; Nummenmaa et al., 2008; Shamay-Tsoory et al., 2009). This study allowed for a deeper exploration of cognitive empathy which arguably, strengthened our understanding of this single component of empathy in young children. This work also gives a new observational measure of cognitive empathy to add to the limited toolbox of measures that can be used with younger children.

Conclusions

This study has offered a novel coding scheme to measure children's cognitive empathy using story stem narratives. Promising results have been evidenced by the direction in which empathy has shown small positive associations with gender, prosocial behavior, and CU traits, in line with existing literature. Our study suggests that story stem narratives are a useful tool that can tap into children's cognitive empathy and should be developed further. The current novel coding scheme shows the potential of using two story stem tasks—Hurt Knee and Three's a Crowd—to understand children's internal worlds within two emotionally charged situations. Children's behavioral and verbal responses to each story scenario demonstrates that the playful nature of this narrative, doll-play approach is developmentally appropriate and presents a unique opportunity to assess cognitive empathy in an authentic way. Specifically, children showed key behaviors of care/concern, helping, and perspective-taking within each of the two tasks. These behaviors indicate some understanding of other's feelings and needs, thus, requiring cognitive empathy. Given the persistent challenges in measuring empathy in young children, there is a continued need for innovative and developmentally appropriate approaches that can capture this key construct. Particularly, with empathy's positive role in social and emotional development, understanding and promoting this behavior is essential for sustaining societies in which people value, support and care for one another. Overall, this study adds value by contributing to empathy assessment, which can benefit our understanding of young children's behavior in a clinical, psychological, and educational capacity.

Data availability statement

The data are not publicly available due to privacy and ethical restrictions, however, analytic codes for this study are available in the public repository, GitHub, at: <https://github.com/Dkwan13/Measuring-Cognitive-Empathy-using-Story-Stem-Narratives>.

Ethics statement

The studies involving humans were approved by Riverside Research Ethics Committee (REC ref:

14/LO/2071). 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

DK: Formal analysis, Methodology, Conceptualization, Writing – original draft, Writing – review & editing. BB: Conceptualization, Project administration, Writing – review & editing, Methodology, Supervision. CO'F: Project administration, Writing – review & editing, Funding acquisition, Data curation. PR: Conceptualization, Funding acquisition, Supervision, Writing – review & editing, Methodology.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study/project was funded by the NIHR Health Technology Assessment NIHR132896.

Acknowledgments

We are grateful to the families and teachers who participated in this follow up study for their generosity in contributing over such a long period. We thank the study PPI group of children and parents for their advice on study design and delivery. We are grateful to the wider Healthy Start Happy Start team and the excellent Study Steering Committee [Kapil Sayal (chair), Samantha Cartwright-Hatton, Florence de Sousa, Richard Emsley, Prasanna Rangadurai and Mara Violato]. CO'F was supported by a UKRI Future Leaders Fellowship [MR/V025686/1].

References

- Ainsworth, M. D. (1970). Attachment, exploration, and separation: illustrated by the behavior of one-year-olds in a strange situation. *Child Dev.* 41, 49–67. doi: 10.2307/1127388
- Arseneault, L., Kim-Cohen, J., Taylor, A., Caspi, A., and Moffitt, T. E. (2005). Psychometric evaluation of 5- and 7-year-old children's self-reports of conduct problems. *J. Abnorm. Child Psychol.* 33, 537–550. doi: 10.1007/s10802-005-6736-5
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends Cogn. Sci.* 6, 248–254. doi: 10.1016/S1364-6613(02)01904-6
- Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* 34, 163–175. doi: 10.1023/B:JADD.0000022607.19833.00
- Basay, Ö., Ezber, S. N., Utku Emre, I.N.C.I., Öztürk, M., Soyuguzel, M. O., and Basay, B. K. (2021). The relation of empathy levels with internalizing and externalizing problems among children and adolescents who refer to child psychiatry outpatients. *Relation* 9, 91–100. doi: 10.4274/nkmj.galenos.2021.820118
- Batson, C. D. (2009). "These things called empathy: eight related but distinct phenomena," in *The Social Neuroscience of Empathy*, eds. J. Decety, and W. Ickes (Cambridge, MA: MIT Press), 3–15.
- Beadle, J. N., Brown, V., Keady, B., Tranel, D., and Paradiso, S. (2012). Trait empathy as a predictor of individual differences in perceived loneliness. *Psychol. Rep.* 110, 3–15. doi: 10.2466/07.09.20.PR0.110.1.3-15
- Belacchi, C., and Farina, E. (2012). Feeling and thinking of others: affective and cognitive empathy and emotion comprehension in prosocial/hostile preschoolers. *Aggress. Behav.* 38, 150–165. doi: 10.1002/ab.21415
- Bensalah, L., Stefaniak, N., Carre, A., and Besche-Richard, C. (2016). The basic empathy scale adapted to french middle childhood: structure and development of empathy. *Behav. Res. Methods* 48, 1410–1420. doi: 10.3758/s13428-015-0650-8
- Bettmann, J. E., and Lundahl, B. W. (2007). Tell me a story: a review of narrative assessments for preschoolers. *Child Adolesc. Soc. Work J.* 24, 455–475. doi: 10.1007/s10560-007-0095-8
- Bowlby, J. (1940). The influence of early environment in the development of neurosis and neurotic character. *Int. J. Psycho-Anal.* 21, 154–178.

Conflict of interest

The author(s) declared that this work 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) declared that generative AI was not 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

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.

Author disclaimer

The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fdpys.2026.1738605/full#supplementary-material>

- Bowlby, J. (1944). Forty-four juvenile thieves: their characters and home-life. *Int. J. Psycho-Anal.* 25, 19–53.
- Bowlby, J. (1958). The nature of the child's tie to his mother. *Int. J. Psycho-Anal.* 39, 350–371.
- Bretherton, I., and Oppenheim, D. (2003). "The MacArthur story stem battery: development, administration, reliability, validity, and reflections about meaning" in *Revealing the Inner Worlds of Young Children: The MacArthur Story Stem Battery and Parent-Child Narratives*, vol. 55 (Oxford: Oxford University Press), 80.
- Bretherton, I., Oppenheim, D., Buchsbaum, H., and Emde, R. N. (1990). *The MacArthur story stem battery* (Unpublished manuscript). doi: 10.1037/t05279-000
- Brody, L. R. (2013). "On understanding gender differences in the expression of emotion: gender roles, socialization, and language," in *Human Feelings* (Milton Park: Routledge), 87–121.
- Bryant, B. K. (1982). An index of empathy for children and adolescents. *Child Dev.* 53, 413–425. doi: 10.2307/1128984
- Buchsbaum, H.K., and Emde, R.N. (1990). Play narratives in 36-month-old children: early moral development and family relationships. *Psychoanalyt. Study Child* 45, 129–155. doi: 10.1080/00797308.1990.11823514
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46. doi: 10.1177/001316446002000104
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334. doi: 10.1007/BF02310555
- Dadds, M. R., Allen, J. L., McGregor, K., Woolgar, M., Viding, E., and Scott, S. (2014). Callous-unemotional traits in children and mechanisms of impaired eye contact during expressions of love: a treatment target? *J. Child Psychol. Psychiatry* 55, 771–780. doi: 10.1111/jcpp.12155
- Davidov, M., Paz, Y., Roth-Hanania, R., Uzevovsky, F., Orlitsky, T., Mankuta, D., and Zahn-Waxler, C. (2021). Caring babies: concern for others in distress during infancy. *Dev. Sci.* 24:e13016. doi: 10.1111/desc.13016
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *Catalog Sel. Doc. Psychol.* 10, 85.
- 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
- Decety, J., and Holvoet, C. (2021). The development of empathy in infancy. *L'Année psychologique* 121, 239–273. doi: 10.3917/anpsy1.213.0239
- Demetriou, H., and Nicholl, B. (2022). Empathy is the mother of invention: emotion and cognition for creativity in the classroom. *Improv. Sch.* 25, 4–21. doi: 10.1177/1365480221989500
- Deneault, A. A., Hammond, S. I., and Madigan, S. (2023). A meta-analysis of child-parent attachment in early childhood and prosociality. *Dev. Psychol.* 59:236. doi: 10.1037/dev0001484
- Denham, S.A., McKinley, M., Couchoud, E. A., and Holt, R. (1990). Emotional and behavioral predictors of preschool peer ratings. *Child Dev.* 61, 1145–1152. doi: 10.2307/1130882
- Devine, R. T., and Hughes, C. (2016). Measuring theory of mind across middle childhood: reliability and validity of the silent films and strange stories tasks. *J. Exp. Child Psychol.* 149, 23–40. doi: 10.1016/j.jecp.2015.07.011
- Devine, R. T., White, N., Ensor, R., and Hughes, C. (2016). Theory of mind in middle childhood: longitudinal associations with executive function and social competence. *Dev. Psychol.* 52:758. doi: 10.1037/dev0000105
- Dunfield, K. A. (2014). A construct divided: prosocial behavior as helping, sharing, and comforting subtypes. *Front. Psychol.* 5:958. doi: 10.3389/fpsyg.2014.00958
- Eisenberg, N., and Fabes, R.A. (1990). Empathy: conceptualization, measurement, and relation to prosocial behavior. *Motiv. Emot.* 14, 131–149. doi: 10.1007/BF00991640
- Eisenberg, N., Fabes, R. A., and Spinrad, T. L. (2006). "Prosocial development," in *Handbook of Child Psychology: Vol. 3: Social, Emotional and Personality Development (6th ed.)*, eds. W. Damon, R. M. Lerner (Eds.), and N. Eisenberg (Vol. Ed.) (Hoboken, NJ: Wiley), 646–718.
- Eisenberg, N., and Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychol. Bull.* 94, 100–131. doi: 10.1037/0033-2909.94.1.100
- Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull.* 101:91. doi: 10.1037/0033-2909.101.1.91
- Eisenberg, N., and Mussen, P. H. (1989). *The Roots of Prosocial Behavior in Children* (Cambridge: Cambridge University Press).
- Eisenberg, N., Spinrad, T. L., and Knafo-Noam, A. (2015). "Prosocial development," in *Handbook of Child Psychology, Vol. 3, Social, Emotional, and Personality Development*, eds. M. E. Lamb, C. G. Coll, and R. M. Lerner, 7th ed. (New York: NYC: Wiley), 610–656.
- Eisenberg-Berg, N. (1979). Development of children's prosocial moral judgement. *Dev. Psychol.* 15, 128–137. doi: 10.1037/0012-1649.15.2.128
- Eisenberg-Berg, N., and Lennon, R. (1980). Altruism and the assessment of empathy in the preschool years. *Child Dev.* 51, 552–557. doi: 10.2307/1129290
- Emde, R. N., Plomin, R., Robinson, J., Corley, R., DeFries, J., Fulker, D. W., and Zahn-Waxler, C. (1992). Temperament, emotion, and cognition at fourteen months: the MacArthur longitudinal twin study. *Child Dev.* 63, 1437–1455. doi: 10.2307/1131567
- Feshbach, N.D., and Roe, K. (1968). Empathy in six- and seven-year olds. *Child Dev.* 34, 133–145. doi: 10.2307/1127365
- Frick, P. J. (2004). *The inventory of callous-unemotional traits* (Unpublished rating scale). The University of New Orleans, New Orleans.
- Frick, P. J., O'Brien, B. S., Wootton, J. M., and McBurnett, K. (1994). Psychopathy and conduct problems in children. *J. Abnorm. Psychol.* 103:700. doi: 10.1037/0021-843X.103.4.700
- Frick, P. J., Ray, J. V., Thornton, L. C., and Kahn, R. E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychol. Bull.* 140, 1–57. doi: 10.1037/a0033076
- Georgiou, G., Kimonis, E. R., and Fanti, K. A. (2019). What do others feel? Cognitive empathy deficits explain the association between callous-unemotional traits and conduct problems among preschool children. *Eur. J. Dev. Psychol.* 16, 633–653. doi: 10.1080/17405629.2018.1478810
- Gerdes, K.E., Segal, E.A., and Lietz, C.A. (2010). Conceptualising and measuring empathy. *Br. J. Soc. Work* 40, 2326–2343. doi: 10.1093/bjsw/bcq048
- Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *J. Child Psychol. Psychiat.* 38, 581–586. doi: 10.1111/j.1469-7610.1997.tb01545.x
- Goodman, R., and Scott, S. (1999). Comparing the strengths and difficulties questionnaire and the child behavior checklist: is small beautiful? *J. Abnorm. Child Psychol.* 27, 17–24. doi: 10.1023/A:1022658222914
- Hay, D. F., and Cook, K. V. (2007). "The transformation of prosocial behavior from infancy to childhood," in *Socioemotional Development in the Toddler Years: Transitions and Transformations*, eds. C. A. Brownell, and C. B. Kopp (New York, NY: Guilford), 100–1310
- Hoffman, M. L. (1990). Empathy and justice motivation. *Motiv. Emot.* 14, 151–172. doi: 10.1007/BF00991641
- Hoffman, M. L. (2000). *Empathy and Moral Development: Implications for Caring and Justice*. New York, NY: Cambridge University Press.
- Hsee, C. K., Hatfield, E., Carlson, J. G., and Chemtob, C. (1990). The effect of power on susceptibility to emotional contagion. *Cogn. Emot.* 4: 327–340. doi: 10.1080/02699939008408081
- Hughes, C., Jaffee, S. R., Happé, F., Taylor, A., Caspi, A., and Moffitt, T. E. (2005). Origins of individual differences in theory of mind: from nature to nurture? *Child Dev.* 76, 356–370. doi: 10.1111/j.1467-8624.2005.00850_a.x
- Iannotti, R. J. (1985). Naturalistic and structural assessments of prosocial behavior in preschool children: the influence of empathy and perspective taking. *Dev. Psychol.* 21, 46–55. doi: 10.1037/0012-1649.21.1.46
- Jolliffe, D., and Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *J. Adolesc.* 29, 589–611. doi: 10.1016/j.adolescence.2005.08.010
- Jolliffe, D., and Farrington, D. P. (2011). Is low empathy related to bullying after controlling for individual and social background variables? *J. Adolesc.* 34, 59–71. doi: 10.1016/j.adolescence.2010.02.001
- Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., and Rhee, S. H. (2008). The developmental origins of a disposition toward empathy: genetic and environmental contributions. *Emotion* 8, 737–752. doi: 10.1037/a0014179
- Kobak, R. (1999). "The emotional dynamics of disruptions in attachment relationships: Implications for theory, research, and clinical intervention," in *Handbook of Attachment: Theory, Research, and Clinical Applications*, eds. J. Cassidy and P. R. Shaver (New York: Guilford Press), 21–43.
- Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., and Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: mixing and matching contexts and perspectives. *Am. J. Psychiatry* 160, 1566–1577. doi: 10.1176/appi.ajp.160.9.1566
- Lennon, R., Eisenberg, N., and Carroll, J. (1983). The assessment of empathy in early childhood. *J. Appl. Dev. Psychol.* 4, 295–302. doi: 10.1016/0193-3973(83)90024-2
- Lilienfeld, S. O. (2003). Comorbidity between and within childhood externalizing and internalizing disorders: reflections and directions. *J. Abnorm. Child Psychol.* 31, 285–291. doi: 10.1023/A:1023229529866
- Lui, J. H., Barry, C. T., and Sacco, D. F. (2016). Callous-unemotional traits and empathy deficits: mediating effects of affective perspective-taking and facial emotion recognition. *Cogn. Emot.* 30, 1049–1062. doi: 10.1080/02699931.2015.1047327
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem Med.* 22, 276–282. doi: 10.11613/BM.2012.031
- Mehrabian, A., and Epstein, N. (1972). A measure of emotional empathy. *J. Pers.* 40, 525–543. doi: 10.1111/j.1467-6494.1972.tb00078.x

- Mieloo, C., Raat, H., van Oort, F., Bevaart, F., Vogel, I., Donker, M., and Jansen, W. (2012). Validity and reliability of the strengths and difficulties questionnaire in 5–6 year olds: differences by gender or by parental education? *PLoS ONE* 7:e36805. doi: 10.1371/journal.pone.0036805
- Miller, P. A., Eisenberg, N., Fabes, R. A., and Shell, R. (1996). Relations of moral reasoning and vicarious emotion to young children's prosocial behavior toward peers and adults. *Dev. Psychol.* 32, 210–219. doi: 10.1037/0012-1649.32.2.210
- Moreno, A. J., Klute, M. M., and Robinson, J. L. (2008). Relational and individual resources as predictors of empathy in early childhood. *Soc. Dev.* 17, 613–637. doi: 10.1111/j.1467-9507.2007.00441.x
- Murray, A. L., Speyer, L. G., Hall, H. A., Valdebenito, S., and Hughes, C. (2021). Teacher versus parent informant measurement invariance of the strengths and difficulties questionnaire. *J. Pediatr. Psychol.* 46, 1249–1257. doi: 10.1093/jpepsy/jsab062
- Neumann, D. L., Chan, R. C., Boyle, G. J., Wang, Y., and Westbury, H. R. (2015). “Measures of empathy: self-report, behavioral, and neuroscientific approaches,” in *Measures of Personality and Social Psychological Constructs* (Cambridge, MA: Academic Press), 257–289. doi: 10.1016/B978-0-12-386915-9.00010-3
- Nummenmaa, L., Hirvonen, J., Parkkola, R., and Hietanen, J. K. (2008). Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. *Neuroimage* 43, 571–580. doi: 10.1016/j.neuroimage.2008.08.014
- Nunnally, J. C. (1967). *Psychometric Theory*. New York: McGraw-Hill.
- Nunnally, J. C. (1978). *Psychometric Theory (2nd ed.)*. New York: McGraw-Hill.
- Nunnally, J. C., and Bernstein, I. H. (1994). *Psychometric Theory (3rd ed.)*. New York: McGraw-Hill.
- O'Farrelly, C., Barker, B., Watt, H., Babalis, D., Bakermans-Kranenburg, M., Byford, S., and Ramchandani, P. (2021). A video-feedback parenting intervention to prevent enduring behaviour problems in at-risk children aged 12–36 months: the Healthy Start, Happy Start RCT. *Health Technol. Assess* 25:1. doi: 10.3310/hta25290
- O'Reilly, J., and Peterson, C. C. (2015). Maltreatment and advanced theory of mind development in school-aged children. *J. Fam. Violence* 30, 93–102. doi: 10.1007/s10896-014-9647-9
- Pederson, R. (2009). 'Empirical research on empathy in medicine: a critical review'. *Patient Educ. Couns.* 76, 307–322. doi: 10.1016/j.pec.2009.06.012
- Petrowski, K., Herold, U., Joraschky, P., von Wyl, A., and Cierpka, M. (2009). The specificity and the development of social-emotional competence in a multi-ethnic-classroom. *Child Adolesc. Psychiatry Ment. Health* 3, 1–10. doi: 10.1186/1753-2000-3-16
- Petrowski, K., Karcz, A., Juen, F., and Cierpka, M. (2014). The ethnic specificity of mental representation and social emotional competence in children. *Mental Health Prevent.* 2, 58–65. doi: 10.1016/j.mhp.2014.11.003
- Qiu, X., Gao, M., Zhu, H., Li, W., and Jiang, R. (2024). Theory of mind, empathy, and prosocial behavior in children and adolescent: a meta-analysis. *Curr. Psychol.* 43, 19690–19707. doi: 10.1007/s12144-024-05762-7
- Reid, C., Davis, H., Horlin, C., Anderson, M., Baughman, N., and Campbell, C. (2013). The Kids' Empathic Development Scale (KEDS): a multi-dimensional measure of empathy in primary school-aged children. *Br. J. Dev. Psychol.* 31, 231–256. doi: 10.1111/bjdp.12002
- Rhee, S. H., Friedman, N. P., Boeldt, D. L., Corley, R. P., Hewitt, J. K., Knafo, A., and Zahn-Waxler, C. (2013). Early concern and disregard for others as predictors of antisocial behavior. *J. Child Psychol. Psychiat.* 54, 157–166. doi: 10.1111/j.1469-7610.2012.02574.x
- Robinson, J., Mantz-Simmons, L., Macfie, J., and the MacArthur Narrative Working Group (1992). *The MacArthur Narrative Coding System* (Unpublished document). University of Colorado Health Sciences Center, Denver.
- Roth-Hanania, R., Davidov, M., and Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behav. Dev.* 34, 447–458. doi: 10.1016/j.infbeh.2011.04.007
- Saarni, C. (1990). “Emotional competence: how emotions and relationships become integrated,” in *Socioemotional Development*, ed. R. A. Thompson (Lincoln: University of Nebraska Press), 115–182.
- Schwartz, S.H., and Howard, J.A. (1984). “Internalized values as motivators of altruism,” in *The Development and Maintenance of Prosocial Behavior: International Perspectives on Positive Development*, eds. E. Staub, D. Bar-Tal, J. Karylowski, and J. Reykowski (New York: Plenum Press), 229–255.
- Sesso, G., Brancati, G. E., Fantozzi, P., Inguaggiato, E., Milone, A., and Masi, G. (2021). Measures of empathy in children and adolescents: a systematic review of questionnaires. *World J Psychiatry* 11:876. doi: 10.5498/wjp.v11.i10.876
- Shamay-Tsoory, S. G., Aharon-Peretz, J., and Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627. doi: 10.1093/brain/awn279
- Simon, P., and Nader-Grosbois, N. (2023). Empathy in preschoolers: exploring profiles and age-and gender-related differences. *Children* 10:1869. doi: 10.3390/children10121869
- Soliman, D., Frydenberg, E., Liang, R., and Deans, J. (2021). Enhancing empathy in preschoolers: a comparison of social and emotional learning approaches. *Educ. Dev. Psychol.* 38, 64–76. doi: 10.1080/20590776.2020.1839883
- Spinrad, T. L., and Eisenberg, N. (2014). “Empathy, prosocial behavior, and positive development in schools,” in *Handbook of Positive Psychology in Schools* (Milton Park: Routledge), 82–98.
- Strauss, C. (2004). Is empathy gendered and, if so, why? An approach from feminist psychological anthropology. *Ethos* 32, 432–457. doi: 10.1525/eth.2004.32.4.432
- Sultan, M. A., and Khan, N. N. (2025). Rethinking empathy development in childhood and adolescence: a call for global, culturally adaptive strategies. *Front. Psychol.* 16:1575249. doi: 10.3389/fpsyg.2025.1575249
- Takahashi, Y., Pease, C. R., Pingault, J.-B., and Viding, E. (2021). Genetic and environmental influences on the developmental trajectory of callous-unemotional traits from childhood to adolescence. *J. Child Psychol. Psychiatry* 62, 414–423. doi: 10.1111/jcpp.13259
- Trimmer, E., McDonald, S., and Rushby, J. A. (2017). Not knowing what I feel: emotional empathy in autism spectrum disorders. *Autism* 21, 450–457. doi: 10.1177/1362361316648520
- Tully, E.C., Ames, A.M., Garcia, S.E., Donohue, M.R. (2016). Quadratic associations between empathy and depression as moderated by emotion dysregulation. *J. Psychol.* 150, 15–35. doi: 10.1080/00223980.2014.992382
- Vreden, C., Buryn-Weitzel, J. C., Atim, S., Donnellan, E., Hoffman, M., Holden, E., and Clay, Z. (2025). Early empathy development: concern and comforting in 9- and 18-month-old infants from Uganda and the UK. *PLoS ONE* 20:e0320371. doi: 10.1371/journal.pone.0320371
- Waller, R., Wagner, N. J., Barstead, M. G., Subar, A., Petersen, J. L., Hyde, J. S., and Hyde, L. W. (2020). A meta-analysis of the associations between callous-unemotional traits and empathy, prosociality, and guilt. *Clin. Psychol. Rev.* 75:101809. doi: 10.1016/j.cpr.2019.101809
- Wang, Z., and Wang, L. (2015). The mind and heart of the social child: developing the empathy and theory of mind scale. *Child Dev. Res.* 2015:171304. doi: 10.1155/2015/171304
- Weiner, S. J., and Auster, S. (2008). From empathy to caring: defining the ideal approach to a healing relationship. *Yale J. Biol. Med.* 80:123.
- Woolgar, M., and Murray, L. (2010). The representation of fathers by children of depressed mothers: refining the meaning of parentification in high-risk samples. *J. Child Psychol. Psychiatry* 51, 621–629. doi: 10.1111/j.1469-7610.2009.02132.x
- Yuval-Adler, S., and Oppenheim, D. (2014). “Story completion play narrative methods for preschool children,” in *Handbook of Research Methods in Early Childhood Education*, Vol. 2 (Charlotte, NC: Information Age Publishing), 323–381.
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., and Chapman, M. (1992). Development of concern for others. *Dev. Psychol.* 28, 126–136. doi: 10.1037/0012-1649.28.1.126
- Zajdel, R. T., Bloom, J. M., Fireman, G., and Larsen, J. T. (2013). Children's understanding and experience of mixed emotions: the roles of age, gender, and empathy. *J. Genet. Psychol.* 174, 582–603. doi: 10.1080/00221325.2012.732125
- Zhou, Q., Valiente, C., and Eisenberg, N. (2003). “Empathy and its measurement,” in *Positive Psychological Assessment: A Handbook of Models and Measures*, eds. S. J. Lopez and C. R. Snyder (Washington, DC: American Psychological Association), 269–284.