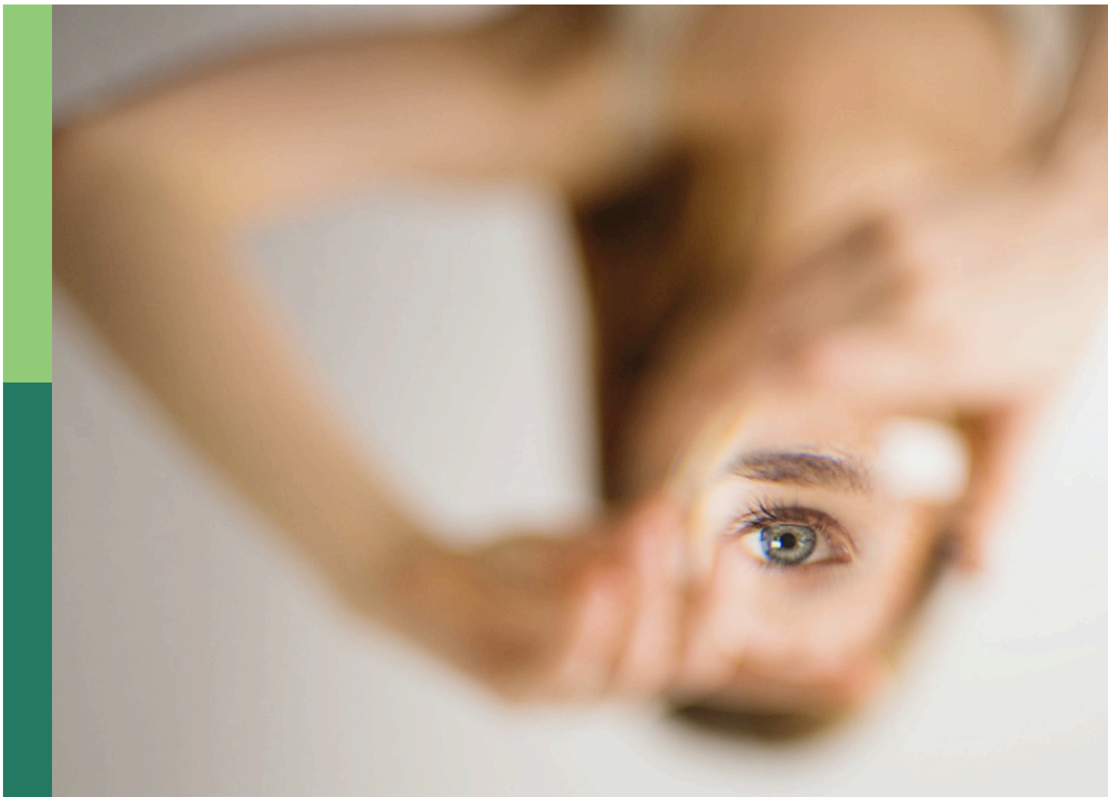
Research Topics

Theory of mind in relation to other cognitive abilities, volume II

Edited by
Ann Dowker and Douglas Frye

Published in
Frontiers in Psychology



ISSN 1664-8714
ISBN 978-2-8325-6542-1
DOI 10.3389/978-2-8325-6542-1

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

Theory of mind in relation to other cognitive abilities, volume II

Topic editors
Ann Dowker — University of Oxford, United Kingdom
Douglas Frye — University of Pennsylvania, United States

Citation
Dowker, A., Frye, D., eds. (2025). *Theory of mind in relation to other cognitive abilities, volume II*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6542-1

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 19 June 2025
DOI 10.3389/fpsyg.2025.1564067

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Eddy J. Davelaar,
Birkbeck, University of London,
United Kingdom

*CORRESPONDENCE
Ann Dowker
ann.dowker@psy.ox.ac.uk
Douglas Frye
dfrye@upenn.edu

RECEIVED 21 January 2025
ACCEPTED 04 June 2025
PUBLISHED 19 June 2025

CITATION
Dowker A and Frye D (2025) Editorial: Theory of mind in relation to other cognitive abilities, volume II. *Front. Psychol.* 16:1564067.
doi: 10.3389/fpsyg.2025.1564067

Editorial: Theory of mind in relation to other cognitive abilities, volume II

Ann Dowker[1]* and Douglas Frye[2]*

[1]Department of Experimental Psychology, Oxford University, Oxford, United Kingdom, [2]Department of Psychology, University of Pennsylvania, Philadelphia, PA, United States

KEYWORDS

theory of mind, false belief (FB), child development, developmental disorders, autism, language, executive functions, central coherence

Editorial on the Research Topic
Theory of mind in relation to other cognitive abilities, volume II

Theory of mind and its development have been the subject of much research over the last 40 years. It is generally thought to be very important in cognitive and social development. However, there is still much debate as to how it should be defined and even as to whether it constitutes a single entity. In particular, there is controversy as to the extent to which it should be seen as a specific cognitive module, or rather as dependent on, or mutually developing with, other cognitive abilities and characteristics, such as language, metacognition, executive function, and cognitive and perceptual styles that emphasize gist vs. detail ("strong" vs. "weak" central coherence). It is also possible that the theory of mind itself has several different components that may be related to different degrees, different cognitive abilities and characteristics. Any relations between the theory of mind and other cognitive characteristics may also vary with age, and may differ between typically developing children and those with autism and other atypical conditions.

Gaining a greater understanding of these issues is important for increasing our understanding of the theory of mind itself, the nature of cognitive development, the similarities and differences between typically and atypically developing children, and whether it is possible to devise interventions to improve the theory of mind, either directly or by improving other abilities. Between 2020 and 2023, we edited a Research Topic on the relation between theory of mind and other cognitive abilities. The current Research Topic is the second volume, and its goal is to extend those earlier findings by bringing together new articles on various aspects of the theory of mind and any concurrent and longitudinal relationships with other cognitive abilities and characteristics.

The articles in this Research Topic can be divided into three broad groups: the nature, correlates and predictors of theory of mind in children; the nature, correlates and predictors of theory of mind in adults; and the role of theory of mind difficulties in disorders.

With regard to theory of mind in children, Ünlüer investigated whether preschool (4- and 5-year-old) children's theory of mind skills and peer relationships predicted their subsequent school adjustment. There were indeed significant relationships. Theory of mind significantly predicted school adjustment as a whole positively predicting liking for school and negatively predicting school avoidance, while prosocial and aggressive behavior toward peers specifically predicted liking for school.

The other articles on children in this Research Topic look at characteristics that predict and may contribute to theory of mind, rather than those that follow on from it. Both articles suggest that certain language skills are important predictors.

de Villiers and de Villiers studied 258 children aged between 3 and 5 over a few months and tested them on three occasions on false belief reasoning and on the possible contributing factors of general language development, complement syntax, vocabulary, and inhibitory control. Cross-sectional and longitudinal regressions showed that all these factors contributed significantly to false belief reasoning. However, by the third assessment, the major proximal contribution was the child's comprehension of syntactic complements. The authors concluded that, as suggested by their earlier training studies, complement syntax makes an important specific contribution to false belief reasoning, but that vocabulary and executive function skills also form pathways to it.

Li and Leung assessed the language skills, executive functions and first-order and second-order false belief reasoning in 150 Mandarin-speaking preschoolers and early primary school children. They found that language was a significant independent predictor of both first-order and second-order false belief reasoning. Executive function predicted first-order false belief reasoning after controlling for age, but not after also controlling for language skills. However, it did continue to be a significant independent predictor of second-order theory of mind even after controlling for both age and language skills.

With regard to theory of mind in adults, one study, similar to Li and Leung's research with children, looked at the possible predictive roles of language and executive function.

Montgomery et al. investigated the theory of mind in adults along with its possible relation to language and executive function. The authors administered a series of advanced theory of mind tasks and tests of vocabulary and executive function to 207 adults. They found that the Strange Stories, Higher-Order False Belief, and Frith-Happé Animation tasks, though relatively weakly correlated, all loaded onto a common factor (?), which they considered to involve perspective-taking, within a narrative context, to represent a protagonist's mental state and use it to predict and explain their actions. This factor was more closely related to vocabulary than to executive function.

Mayrand et al. carried out a rather different type of study, looking at adults' interpretation of information communicated by gaze. They investigated how spatially dissociated vs. spatially combined effects of gaze (i.e., cases where an agent's inferred mental content implied by gaze is discrepant with the directional information communicated by gaze, vs. cases where the two types of cues provide concordant information) influence participants' target performance. They found that performance was worse when cue direction and mental content were discordant than when they were concordant. This effect was more marked when a social avatar served as a cue than when a comparison arrow was the cue. These findings suggest that a typical gaze communicates information about both what a person is attending to and the location of their attention.

Other studies examined the theory of mind in relation to disorders. The disorder that has been studied the most over the years with regard to associated limitations in theory of mind is autism spectrum disorder, as reflected in the articles in this Research Topic.

Qiao et al. conducted a comprehensive review of the literature from the past 30 years on the broader autism phenotype. First, they used the Web of Science Core Collection database to find articles on the autism phenotype in general published between 1994 and 2024. They then used the CiteSpace and VOS viewer software to further visualize and analyze the citations. They identified a total of 1,075 articles related to the broader autism phenotype. The annual number of publications on the subject has increased over the past 30 years. The largest number of publications came from the United States, followed by England and Canada. The United States also ranked first in terms of the extent to which its publications were cited.

Liu used photographs of social scenes to compare adolescents with autism spectrum disorder and controls on their ability to reason and make inferences about people's intentions. The adolescents with ASD performed significantly worse than the controls in making inferences about intentions. However, their ability to make physical causal inferences was unimpaired. Liu also investigated the relations between performance in these tasks and performance in tests of working memory and attention. Among individuals with ASD, working memory predicted physical causal inference, while divided attention predicted inferences about intention.

Gambra et al. investigated central coherence in typically-developing 6- to 11-year-olds and in several clinical groups: children with autism spectrum disorder; children with ADHD alone; children with ADHD and a non-verbal learning disability; and children with a social communication disorder. They used Gambra's (2020) Central Coherence Test to investigate children's use of context to make inferences and solve problems, rather than focusing predominantly on details. Children with ADHD alone did not differ from controls, but all the other clinical groups showed weaker central coherence than the controls. Those with ASD did not differ from the other clinical groups. Thus, autism was not the only condition associated with weak central coherence. The other disorders associated with weak central coherence in this study are also conditions that one might expect to be associated with theory of mind limitations, although this association was not explicitly tested in the present study.

Limitations in theory of mind have been found not only in patients with developmental disorders, but also in deaf children with hearing parents, possibly as a result of limited exposure to conversations about mental states. Wu et al. investigated whether deaf college students could be trained to improve their theory of mind. They trained 40 deaf students in theory of mind and compared them with 40 active controls, who received physical conversation training. The students, who were trained in theory of mind, improved significantly from pre-test to post-test on both cognitive and affective theory of mind tasks, and performed significantly better than the controls at post-test. These results suggest that training can improve theory of mind performance in deaf students.

The findings reported in the current Research Topic give rise to possibilities for further research. For example, to what extent and in what ways might the relations between the theory of mind and other cognitive abilities change with age? Do relations between the theory of mind and other cognitive abilities vary with culture and language? How do relations between the theory of mind and other cognitive abilities differ across different developmental disorders, and what implications might this have for intervention? The entire area of research is important and multifaceted and holds much promise for future development.

Author contributions

AD: Writing – original draft. DF: Writing – original draft.

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

Gambra, L. (2020). *Central Coherence: Design and Validation of an Assessment and Study Test in Procedural Learning Disorder*. Pamplona: University of Navarra.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 17 April 2024
DOI 10.3389/fpsyg.2024.1361878

OPEN ACCESS

EDITED BY
Douglas Frye,
University of Pennsylvania, United States

REVIEWED BY
Elena Cavallini,
University of Pavia, Italy
Elisabetta Lombardi,
eCampus University, Italy

*CORRESPONDENCE
Yang Wu
1508549839@qq.com

RECEIVED 27 December 2023
ACCEPTED 08 April 2024
PUBLISHED 17 April 2024

CITATION
Wu Y, Liu X and Zhang S (2024) Training deaf college students to improve their theory of mind: based on a two-component model. *Front. Psychol.* 15:1361878.
doi: 10.3389/fpsyg.2024.1361878

Training deaf college students to improve their theory of mind: based on a two-component model

Yang Wu[1]*, Xiping Liu[2] and Shengnan Zhang[3]

[1]School of Marxism, Tianjin University of Technology, Tianjin, China, [2]Faculty of Psychology, Tianjin Normal University, Tianjin, China, [3]Faculty of Education, Henan Normal University, Xinxiang, China

This paper explored the training methods to improve the level of deaf college students' ToM. Eighty deaf college students were selected as participants and randomly divided into experimental group and control group. The ToM training group received ToM training; The non-ToM training group received physical-conversation training. Cognitive ToM task and affective ToM task were used to investigate the training effect. After training, the level of ToM of deaf college students who received ToM training was significantly improved. The results show that ToM training can effectively promote the level of deaf college students' ToM.

KEYWORDS

training, deaf Chinese college students, theory of mind, cognitive psychology, development psychology

Introduction

Theory of Mind (ToM) is the ability to make inferences about the psychological states of others and predict or explain their behavior with reference to their mental states, feelings, beliefs, and desires (Premack and Woodruff, 1978; Wellman and Liu, 2004). ToM is a key factor in the development of interpersonal skills and the formation and maintenance of social relationships (Devine and Hughes, 2016; Lecciso et al., 2016; Petrocchi et al., 2021).

Studies of deaf preschool children and deaf school-aged children have found that ToM delay can persist, at least in some respects, until the age of 12 or 13 years (O'Reilly et al., 2014). In a study of social perception, Lecciso et al. (2013) measured this component of ToM using the Eye Test task (Baron-Cohen et al., 2001) and found that deaf children, aged 5–14 years, performed worse than children with normal hearing. Although ToM of deaf individuals with deaf parents may equal that of hearing individuals later in development, evidence of delayed development after adulthood remains for deaf individuals without deaf parents (O'Reilly et al., 2014). A study evaluated deaf college students' understanding of sarcasm, second-order false beliefs, and double deception. The results showed that these individuals scored significantly lower on all three ToM tests than their hearing peers. The same results were found in Edwards et al. (2021), where deaf college students performed worse than their hearing peers on sarcasm, metaphor comprehension tasks, and reasoning tasks.

The reason for this may be that deaf people with hearing parents experience less conversation at home than deaf children with deaf parents (Peterson and Slaughter, 2006). Parents' conversations with deaf children lack high-level facilitative discourse, such as mental state conversation, and are more indicative than those with hearing children (Ambrose et al., 2015). This situation results in deaf children often having no one at home with whom they can freely talk about observed thoughts, feelings, and other mental states. This development delay indirectly affects deaf college students' social ability and the formation of social relationships. Studies have shown that a decline in the ability to understand mental states may be positively correlated with a decline in social activities (Bailey et al., 2008), and negatively correlated with the scale of intimate social networks (Radecki et al., 2019). Deaf college students who are about to enter the society are at a special stage. They will soon independently bear social and personal responsibility; therefore, they often demonstrate a need improve their level of ToM and social adaptability during this critical transition. Therefore, it is very important to find ways to quickly improve their theory of mind. In view of the previous studies, ToM can be divided into cognitive ToM and affective ToM, so this study hopes to find a training method that can improve the cognitive ToM and affective ToM of deaf college students at the same time, so as to help deaf college students improve their social adaptability.

Factors influencing training

Thus far, studies promoting ToM levels have typically been conducted on preschoolers, school-age children, individuals with autism, and older individuals (Wellman et al., 2001; Caputi et al., 2021). Training studies in older adults have shown that ToM training based on mental state conversations is effective in improving social cognitive skills in nursing home residents. Older people are able to transform the skills acquired during training into new materials, thereby improving their quality of life (Cavallini et al., 2021). Such training also shows the importance of mental state conversations in promoting ToM levels.

In addition to the mental state conversations, there are other details that can be focused upon when training. The first is whether reading materials with a large number of mental state words are included in the training process. In adults, Kidd and Castano (2013) demonstrated the effect of reading literary fiction on ToM skill development. People who read literary fiction performed better on a ToM task than those who read non-fiction. This association may be a function the large number of words related to ToM typically encountered in novels. In addition, the researchers found that mothers' use of mental state words while reading picture books was strongly correlated with their children's performance on false belief tests, and this association lasted for at least 1 year (Adrian et al., 2007). In the process of reading, the reader is required to make inferences about the characters and the plot of the story, and to understand the vocabulary of mental states used in the story, all of which support ToM development (De Mulder et al., 2017).

The second training detail is whether there is a large amount of interpretation in the training process. Research on ToM training shows that informing a child whether their answer is correct and explaining why it is necessary can make a difference in ToM ability (Clements et al., 2000; Melot and Angeard, 2003). Giving feedback and explanation in conversations can make a connection between a character's inner state and outward behavior (Slaughter and Peterson, 2012). For example, Schult and Wellman (1997) found that preschool children explain more than 85% of voluntary behavior and mistakes made by others through their mental states, and that explanation can help them build their personal worldview.

The final training detail is a focus on facial cues. While individuals without hearing loss rely heavily on facial changes as supplementary cues to comprehend conversational content, for deaf individuals, this is a crucial factor in communication. Furthermore, native sign language places greater emphasis on the eye area (Emmorey et al., 2009), while late sign language learners tend to focus more on the mouth area (Muir and Richardson, 2002). Therefore, facilitating easier observation of facial cues during training is highly important for deaf individuals.

Current study design

ToM training utilized in this study was based on the comprehensive ToM treatment approach developed by Cavallini for older adults. First, during the training process, participants engaged in group discussions. Participants were arranged to sit in a circle, enabling them to visually perceive each other's facial expressions and sign language cues synchronously. It is worth noting that most sign languages have distinct mouth shapes associated with gesture movement and timing (Boyes Braem and Sutton-Spence, 2001; Sutton-Spence, 2007). Second, an extensive collection of reading materials containing mental state vocabulary was prepared for the training sessions. These materials encompassed ToM state verbs which are particularly focused on abstract internal states and cognitive processes. A study by de Villiers and de Villiers (2014) found that employing mental state verbs directly distinguishes mental content from reality and plays a central role in understanding false beliefs. Finally, ample explanations were incorporated into the training. Interpretation serves as a fundamental aspect of cognition and occurs naturally within everyday social interactions (Keil, 2006). In discussions revolving around mental states, feedback and interpretation appear to be crucial elements (Clements et al., 2000). Throughout this study's training process, participants were required to explain trainers' questions, as well as address any doubts, while providing their own answers. Through this interpretive process, we hypothesized that participants would enhance their self-understanding and comprehension of the world, including ToM concepts.

In addition, previous research designs have used different ToM stories to examine the overall effect of training. These testing methods are unable to examine multidimensional changes in ToM and compare differences between dimensions. The current study attempts to investigate training effects on the cognitive and affective dimensions of ToM based on the two-component model. Some researchers believe that ToM is composed of two parts: cognitive processing and emotional processing (Shamay-Tsoory et al., 2006). Brothers and Ring (1992) believe that cognitive ToM, which refers to the inference of knowledge and belief, develops around 4–5 years old. Affective ToM, which refers to the inference of emotion, is thought to develop around 2–4 years old (MacDonald et al., 1996). Many studies have shown that affective ToM and cognitive ToM can be separated in behavior (Hynes et al., 2006; Völlm et al., 2006; Shamay-Tsoory and Aharon-Peretz, 2007). Extensive studies using brain imaging techniques have simultaneously demonstrated the existence of a partial separation mechanism at the neural level between cognitive and affective ToM (Dvash and Shamay-Tsoory's, 2014). This neuroanatomical separation between affective and cognitive regions has also been shown in deaf

10.3389/fpsyg.2024.1361878

individuals (Hao et al., 2010; Lecciso et al., 2016). This study examines the theory of mind level of deaf college students from two component and further verifies the existence of separate component from the results.

Finally, previous training studies have often used test materials containing text to examine training effects, which easily adds additional processing burden to the deaf group in completing the task, possibly reducing the level of ToM assessment. In this study, two picture only tasks were selected, which could be performed without relying on text processing, and the effect of training could be objectively judged.

Therefore, this study takes mental state as the focus of ToM training, and combines the above factors affecting ToM changes to develop suitable training content for deaf college students, in an attempt to improve the ToM level of deaf college students through four training courses. Cognitive and affective ToM levels of deaf college students before and after training were investigated through two experimental procedures, each focusing on cognitive or affective ToM. Two months after the training, cognitive ToM and affective ToM were tested again to investigate the long-term effects of the training.

TABLE 1 Participant characteristics.

	Participant characteristic	ToM training (*n* = 40)		No-ToM training (*n* = 40)	
		M	*SD*	*M*	*SD*
	Age	20.25	1.14	20.05	1.53
	Raven	40.20	3.21	39.75	3.72
Female, *n*		21		19	
Grade, *n*	Freshman	14		13	
	Sophomore	13		14	
	Junior	13		13	
The hearing status, n	Severe hearing loss	10		13	
	Profound hearing loss	26		22	
	Complete or total hearing loss/ deafness	4		5	

Scores in parenthesis refer to standard deviation; Parental education level: 1 = no/primary education; 2 = lower general secondary education; 3 = higher general secondary education; 4 = college/university; Grades of hearing loss is based on the grading system published by WHO in 2021.

Methods

Participants

A total of 80 deaf college students (M = 20.15 years old, SD = 1.34; 50%, female) were recruited through information provided on the website of a special education university in Northern China. Participants were required to be 18 years of age or older and have hearing loss. Participants were also required to pass a university entrance exam that included reading and writing content, and demonstrate a proficient level of written and spoken Chinese or Chinese Sign language. None of the students were clinically diagnosed with any mental disorder. An intelligence test was completed by participants prior to group assignment, and participants did not receive any tangible reward (such as money or gifts). Participants were randomly assigned to two groups: the ToM training group (18–22 years old) and the no-ToM training group (18–22 years old). There was no significant difference in age [$F_{(1, 78)} = 1.60$, $p > 0.05$], grade [$F_{(1, 78)} = 0.35$, $p > 0.05$], sex [$F_{(1, 78)} = 0.13$, $p > 0.05$], or intelligence [$F_{(1, 78)} = 2.50$, $p > 0.05$] between groups. The descriptive statistics of age, grade, intelligence, male to female ratio, and hearing level are shown in Table 1. All participants completed an informed consent form before starting the study.

Measures and procedure

Research design

This study used a 2 (group: ToM training vs. no ToM training) × 3 (timepoint: pre-test, post-test 1, post-test 2) two-factor-mixed experimental design. Group was the between-subject variable and timepoint was the within-subject variable.

Intelligence test

Raven's Reasoning Test is a kind of nonverbal intelligence test, which was compiled by Raven in 1938, and revised by Zhang and Wang (1985), which established Chinese urban norms. The test consists of five series, each consisting of 12 questions, all of which are graphic reasoning questions. The standard version of the test contains 60 items. Each item consists of a large graph and several small graphs. The lower right corner of the large graph is missing. The task of the test subjects is to find out which part is missing in the lower right corner of the large graph. The test is scored on a 2-point scale, meaning correct answers are scored as 1 point and incorrect answers are scored as 0 points. The partial reliability of the test was 0.95, 0.82 for retest at half month intervals, and 0.79 for retest at 1 month intervals, indicating good reliability. The prediction validity is good.

Cognitive ToM

A total of 36 items were presented to participants (18 were intentional reasoning tasks and 18 were physical reasoning tasks) (Murdaugh et al., 2014; Thye et al., 2018). In the intentional reasoning task, participants were shown a comic strip with three images of objects or people and were asked to choose a logical ending from three options. In the physical reasoning task, subjects were presented with a cartoon image depicting a physical behavior, while in the intentional reasoning condition, subjects were presented with three images depicting goal-oriented and social interaction behavior, with specific presentation modes (Figure 1).

Affective ToM

In a revised version of the eye task, participants were presented with the eye area of a face and instructed to choose the most suitable word from a list of four options that best described the mental state conveyed by the character's eyes. In the control condition, participants were instructed to identify the gender status of the individuals depicted in the photos, with 72 pictures, including 36 that depicted mental states and 36 depicted gender status (Baron-Cohen et al., 2001; Thye et al., 2018) (Figure 2).

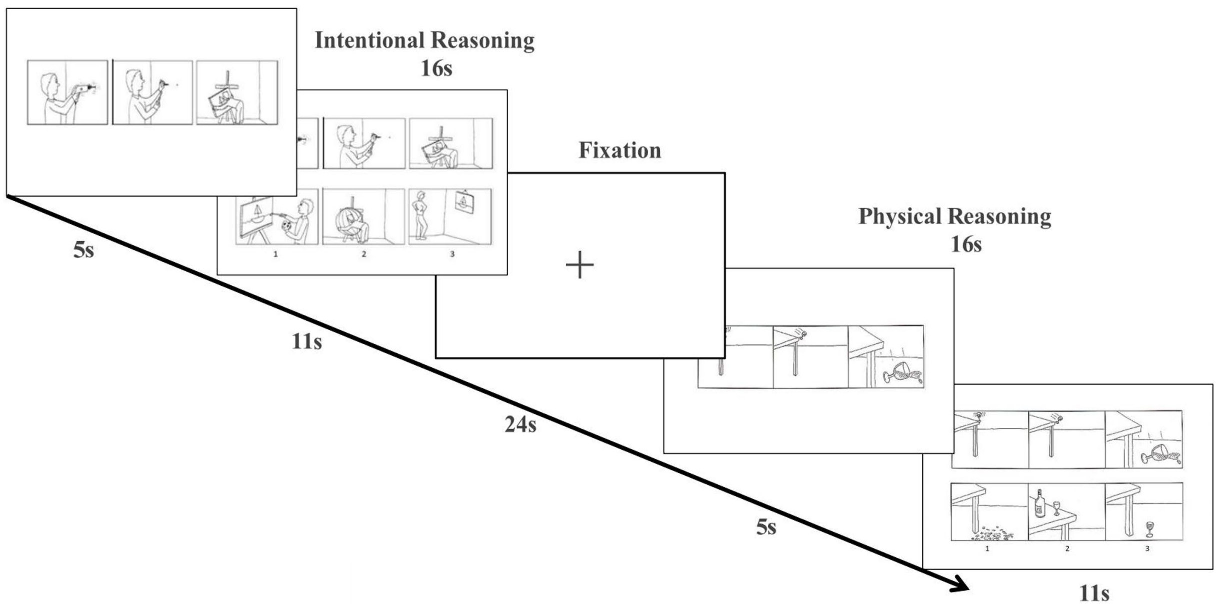


FIGURE 1
Cognitive ToM task. Task design and example stimuli. The correct answers for the example stimuli pictured here are "3" (intentional condition) and "1" (physical condition). s, seconds.

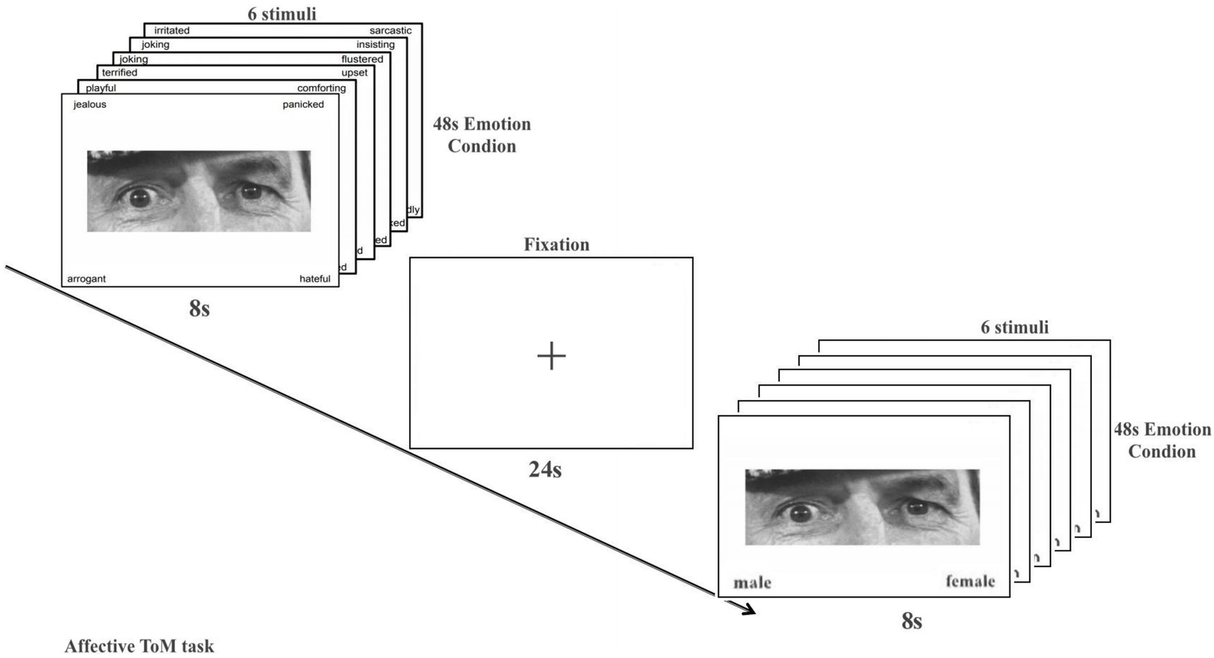


FIGURE 2
Affective ToM task. Task design and example stimuli. The correct answers for the example stimuli pictured here are "panicked" (emotion condition) and "male" (gender condition). s, seconds.

Procedure

Participants were randomly divided into four groups of 20 and trained simultaneously in four classrooms. Two groups received ToM training with conversations focusing on mental states, while other two groups received no-ToM training, with conversations focusing on physical aspects. The training content was adapted from Cavallini's ToM training program developed in Cavallini et al. (2015), specifically tailored for deaf college students by incorporating classroom, dormitory, and community activity scenes. Each training group had a sign language interpreter to facilitate comprehension of the material. Considering variations in hearing impairment among participants, visual materials were used for presentation purposes.

All participants underwent testing at three timepoints: pre-test, post-test 1, and post-test 2; these assessments were administered by individuals who did not participate in any training. Pre-tests included the Raven Intelligence Test, as well as cognitive and affective ToM tasks. There were four training sessions. The duration of each weekly session was 90 min, and took place over the course of 1 month. The ToM training group was exposed to four stories, each containing conversations about mental states, either imagined or real. Participants were asked to answer questions during the training, and discussed their answers, focusing on the changes in the mental state of the characters in the story and the changes in the mental state of the participants themselves. In each lesson, trainers guided participants through exercises aimed at familiarizing them with mental state verbs and distinguishing them from other types of verbs through practice activities involving substitution within the stories themselves. Each time the no-ToM training group was exposed to the same story, with the only change being questions and discussions based on physical characteristics of the characters. The structure and length of the physical characteristics story was the same as that of the story presented to the ToM training group. The post-training tests took place immediately after training and 2 months after the training, and included the cognitive and affective ToM tests.

Analytic plan

We hypothesize that ToM training has positive effects on both intentional reasoning and emotional reasoning in the ToM training group, while non-ToM training has positive effects on both physical reasoning and gender reasoning in the non-ToM training group. SPSS 25.0 (IBM Corp., Armonk, NY) was used for statistical analysis. Descriptive statistical analysis and mixed measure ANOVA were used, respectively, to examine the homogeneity of the two groups on different tasks before and after training and the changes before and after training. If the interaction is significant after repeated measurement analysis, simple effect analysis is used to further analyze the results. If the interaction is not significant, the paired *T*-test is used to further analyze the results.

Results

Table 2 shows the performance of two groups of deaf college students in completing four tasks before and after training. To investigate the effect of training on the performance of the two groups in the cognitive ToM task, a 2 (group: ToM training group, no-ToM training group) × 3 (timepoint: pre-test, post-test 1, post-test 2) mixed measure ANOVA analysis was performed.

Results showed that the accuracy rate (ACC) interaction was significant in intentional reasoning, $F_{(2, 77)} = 6.12$, $p < 0.05$, $\eta^2 = 0.137$. Simple effect analysis showed that there was no significant difference between ToM training group and no-ToM training group for pre-test conditions, $F_{(1, 78)} = 0.017$, $p > 0.05$, $\eta^2 = 0.000$; at the first post-test, the difference between the ToM training and no-ToM training groups was significant, $F_{(1, 78)} = 7.54$, $p < 0.05$, $\eta^2 = 0.088$; at the second post-test, the difference between the ToM and no-ToM training groups was also significant, $F_{(1, 78)} = 36.27$, $p < 0.05$, $\eta^2 = 0.317$. Moreover, there were significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 11.47$, $p < 0.001$, $\eta^2 = 0.23$, but no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 0.014$, $p > 0.05$, $\eta^2 = 0.000$. The results showed that the ACC changes of intentional reasoning were significantly different between the ToM and no-ToM training groups. ToM training had a positive effect on the ACC of intentional reasoning in the ToM training group, and this improvement was maintained until the second post-test, which demonstrates a long-term effect.

The reaction time (RT) interaction was significant, $F_{(2,77)} = 3.89$, $p < 0.05$, $\eta^2 = 0.048$. Simple effect analysis showed that under pre-test conditions, there was no significant difference between the ToM and no-ToM training groups, $F_{(1,78)} = 2.59$, $p > 0.05$, $\eta^2 = 0.032$; at the first post-test, the difference between the ToM and no-ToM training groups was significant, $F_{(1,78)} = 6.17$, $p < 0.05$, $\eta^2 = 0.015$; at the second post-test, the difference between the ToM and no-ToM training groups was also significant, $F_{(1,78)} = 10.04$, $p < 0.05$, $\eta^2 = 0.058$. Moreover, there were significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 4.17$, $p < 0.005$, $\eta^2 = 0.098$, but no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 1.95$, $p > 0.05$, $\eta^2 = 0.048$. Results showed that the RT in the ToM training group became shorter after training, while the RT in the no-ToM training group did not change significantly. ToM training exerted a positive effect on the RT of intentional reasoning in the ToM training group, shortening the RT of intentional reasoning, and this effect was maintained until the second post-test, demonstrating a long-term effect.

In the physical reasoning test, the ACC interaction was no significant, $F_{(2,77)} = 0.657$, $p > 0.05$, $\eta^2 = 0.017$. Simple effect analysis showed that there was no significant difference between the ToM and no-ToM training groups at the pre-test timepoint, $F_{(1,78)} = 0.027$, $p > 0.05$, $\eta^2 = 0.000$; at the first post-test, there was no significant difference between the ToM and no-ToM training groups, $F_{(1,78)} = 0.958$, $p > 0.05$, $\eta^2 = 0.012$; at the second post-test timepoint, there was no significant difference between the ToM and no-ToM training groups, $F_{(1,78)} = 0.074$, $p > 0.05$, $\eta^2 = 0.001$. Moreover, there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 0.34$, $p > 0.005$, $\eta^2 = 0.009$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 0.45$, $p > 0.05$, $\eta^2 = 0.012$. Results showed no significant difference between the two groups measured before and after in the ACC measurement. Physical training had no effect on ACC performance in physical reasoning in the no-ToM training group. This result will be examined in further detail in the discussion section.

RT interaction was not significant, $F_{(2,77)} = 0.834$, $p > 0.05$, $\eta^2 = 0.011$. But paired *t*-test analysis showed that at the pre-test timepoint, there was no significant difference between the ToM and no-ToM training groups, $T_{(39)} = -1.337$, $p > 0.05$; at the first post-test, the difference between the ToM and no-ToM training groups was significant, $T_{(39)} = -3.347$, $p < 0.05$; at the second post-test, the difference between the ToM and no-ToM training groups was also significant, $T_{(39)} = -2.423$, $p < 0.05$. Moreover, there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 11.68$, $p < 0.005$, $\eta^2 = 0.233$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 6.91$, $p < 0.05$, $\eta^2 = 0.152$. In physical reasoning, there were significant differences between the two groups in RT performance at all three testing timepoints. Whether this result indicates that physical training has a positive effect on the no-ToM group requires further discussion. Because in the results of the three tests of the ToM training group, the RT also showed a trend of shorter and shorter.

TABLE 2 Mean values and standard deviations for task performance as a function of Group Factor (ToM and No-ToM groups) and Timepoint (pre, post-1, and post-2).

				Pre		Post-1		Post-2	
				M	*SD*	*M*	*SD*	*M*	*SD*
ToM training	Cognitive ToM	Intentional reasoning	ACC	0.6028	0.13	0.7050	0.11	0.6989	0.13
			RT	4840.1	942.9	4530.8	1027.1	4258.7	948.2
		Physical reasoning	ACC	0.8247	0.14	0.8412	0.14	0.8299	0.14
			RT	4029.6	849.3	3383.3	710.7	3192.1	565.9
	Affective ToM	Emotion reasoning	ACC	0.4757	0.11	0.5417	0.10	0.5348	0.07
			RT	4173.9	557.2	4221.3	552.6	4213.3	556.9
		Gender reasoning	ACC	0.7372	0.10	0.7400	0.04	0.7619	0.08
			RT	2782.6	661.4	2669.2	448.0	2563.1	539.1
No ToM training	Cognitive ToM	Intentional reasoning	ACC	0.6069	0.15	0.6117	0.19	0.6114	0.11
			RT	4494.4	979.3	4266.4	1158.3	4696.4	827.6
		Physical reasoning	ACC	0.8194	0.14	0.8095	0.14	0.8389	0.14
			RT	4261.4	1129.9	4099.2	1111.5	3554.4	784.0
	Affective ToM	Emotion reasoning	ACC	0.5069	0.09	0.5208	0.12	0.5117	0.08
			RT	3891.6	528.2	4147.5	592.87	4019.1	513.2
		Gender reasoning	ACC	0.7380	0.07	0.7675	0.06	0.7823	0.09
			RT	2726.6	473.7	2690.1	507.9	2501.7	489.0

M, mean; SD, standard deviation; ACC, correct rate; RT, reaction time.

To further examine the impact of training on participant performance in the affective ToM Tasks, a 2 (group: ToM training group, no-ToM training group) × 3 (timepoint: pre-test, post-test 1, post-test 2) repeated measures analysis of variance was conducted.

Results showed that the ACC interaction was not significant in emotional reasoning, $F_{(2,77)} = 2.52$, $p > 0.05$, $\eta^2 = 0.062$. But there were significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 6.404$, $p < 0.005$, $\eta^2 = 0.143$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 0.165$, $p > 0.05$, $\eta^2 = 0.004$. But there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 6.404$, $p < 0.005$, $\eta^2 = 0.143$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 0165$, $p > 0.05$, $\eta^2 = 0.004$. On the affective ToM task, there was no significant difference between the two groups on the pair-wise comparison of the three tests of the ACC for emotional reasoning. But the ToM training group showed significant differences across the three measurements. The results showed that ToM training significantly improved the ACC of ToM training group. And this result shows that physical reasoning training had no effect on ACC.

The RT interaction was not significant, $F_{(2,77)} = 0.467$, $p > 0.05$, $\eta^2 = 0.006$. And there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 0.448$, $p > 0.005$, $\eta^2 = 0.012$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 2.009$, $p > 0.05$, $\eta^2 = 0.05$. Moreover, there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 0.448$, $p > 0.05$, $\eta^2 = 0.012$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 2.009$, $p > 0.05$, $\eta^2 = 0.05$. The results showed that RT performance in emotional reasoning did not differ significantly between the two groups on any of the three measures. The results showed that ToM training had no effect on RT of emotional reasoning in the ToM training group. The reason why it could not help the ToM training group shorten RT is worth being discussed.

In gender reasoning, ACC interaction was not significant, $F_{(2, 77)} = 0.418$, $p > 0.05$, $\eta^2 = 0.009$. And there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 0.903$, $p > 0.05$, $\eta^2 = 0.037$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 1.786$, $p > 0.05$, $\eta^2 = 0.071$. Moreover, there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 0.903$, $p > 0.05$, $\eta^2 = 0.037$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 1.786$, $p > 0.05$, $\eta^2 = 0.071$. In gender reasoning, there was no significant difference in ACC between the two groups across all three tests. The ToM training group also did not show significant differences across the three tests as a whole. The results showed that physical training had no effect on the ACC of gender reasoning in the no-ToM training group. No change in the ACC of gender reasoning in the no-ToM training group is worth being discussed.

RT interaction was not significant, $F_{(2, 77)} = 0.114$, $p > 0.05$, $\eta^2 = 0.002$. And there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 1.103$, $p > 0.05$, $\eta^2 = 0.045$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 1.51$, $p > 0.05$, $\eta^2 = 0.06$. Moreover, there were no significant differences in the three measurement results of the ToM training group, $F_{(2, 77)} = 1.103$, $p > 0.05$, $\eta^2 = 0.045$, and no significant differences in the three measurements for the no-ToM training group, $F_{(2, 77)} = 1.51$, $p > 0.05$, $\eta^2 = 0.06$. In gender reasoning, there was no significant difference in RT between the two groups in any of the three comparison tests. The two groups as a whole also did not show significant differences across the three tests. The results showed that physical training had no effect on the RT of gender reasoning in the no-ToM training group. This result will also be discussed in the discussion section.

Discussion

In the current study, a new test method was adopted to assess ToM training efficacy. In previous studies of ToM training, questionnaires were used to determine results (Cavallini et al., 2015; Tucci et al., 2016; Caputi et al., 2021). In these studies, significant duplication between the content of the questionnaire and the content of the training existed; hence, familiarity with the content of the test may have resulted in the illusion of improvement. To address this shortcoming, a detection method that was consistent with the training content, but did not exactly duplicate the content, was needed. Therefore, the detection method in this study included intentional reasoning, emotional reasoning, physical reasoning, and gender reasoning, with no duplication between the test content and the training content. It can be seen from the training results that after ToM training, the scores of the ToM group in ACC of intention reasoning and emotional reasoning have been significantly improved, and the performance is better than that of the non-ToM group, and the task scores have been significantly improved in the post-test, which proves that ToM training is effective in improving ToM.

The key role of mental state conversation

Differences in the language learning experience of deaf people result in a special developmental environment of ToM. Deaf people with hearing parents often grow up in an environment of oral communication, which is limited by their limited oral ability (Lederberg et al., 2013). The limited ability to speak results in individuals within the deaf community having limited opportunities to participate in a variety of conversations, including discussions of different perspectives and the use of terms that describe mental states and a range of other syntactic structures (de Villiers and de Villiers, 2014). Studies have shown that deaf children whose exposure to sign language is delayed have fewer opportunities to talk at home than children who are born deaf and have immediate access to sign language (Peterson and Slaughter, 2006). Success on the false belief task shows a significant delay (Peterson and Siegal, 1999). The reason why an environment that contains conversations about mental states is so important to the development of ToM may be that it draws attention to the inner states of others and shapes their expectations or experiences of interpersonal events (Slaughter and Peterson, 2012). Exposure to environments that contain conversations about mental states not only encourages individuals to pay closer attention to mental states, but can also help individuals learn how to better spot specific mental states, and thus have a better understanding of what is happening in a particular social context (Apperly, 2011). Many studies have shown that mental state conversation is the key factor determining individual differences in ToM (de Rosnay and Hughes, 2006). Mental state conversation can predict ToM development (Lecce et al., 2014). As a result, a large amount of mental state conversation was incorporated into this training, including reading material and interactive content. The deaf college students in the current study were effectively immersed in the environment of mental state conversation in an effort to improve their ToM level. Deaf college students have improved their theory of mind, which is consistent with previous research results (Bianco et al., 2021; Lombardi et al., 2022).

The role of explanation in training

In verbal conversation, explanation is crucial (Melot and Angeard, 2003). Giving explanations in a conversation can help to show the relationship between a character's inner state and outer behavior (Dunn et al., 1991). A study by Marschark, Spencer, Adams, and Sapere showed that hearing adults are more likely to control conversation when a child is hearing impaired, and to avoid confusion, deaf children rarely mention their own doubts, misunderstandings, or incorrect ideas (Marschark et al., 2011). Even children who receive cochlear implants in early childhood are less likely to be exposed to the informal causal explanatory conversations about people associated with early false belief success that occur in families without hearing deficits (Peterson and Slaughter, 2003). Studies that ask deaf children to explain rules of school behavior have found that deaf children perform significantly worse than non-deaf children (Rachford and Furth, 1986). Calderon and Greenber similarly found that deaf children often misunderstand or ignore the "why" question in conversation. Researchers believe that limited interpretation and limited experience deprive many deaf children of legitimate opportunities to learn how to understand others (Calderon and Greenberg, 2003). Taking the above into account, the primary reason for the success of the intervention in the current study is primarily due to the interactions participants had with trainers. The trainer consciously reminded participants to explain their choices, and to explain the mental state and purpose of the behavior of the hero in the story. When the participant demonstrated difficulty in explanation, the trainer invited other participants to assist in the explanation and help to craft a correct explanation. Unfortunately, interpretation was not examined as a separate factor, which will be considered in subsequent studies.

Training effects of training on ToM

Intentional and emotional reasoning was primarily improved in the current study through the analysis, sharing, and explanation of ToM stories. In contrast, the no-ToM training group only practiced physical reasoning. In cognitive ToM, intentional reasoning ability of the ToM training group was improved after training. ACC was increased and RT was decreased.

Interestingly, although the no-ToM training group was trained in physical reasoning, there was no significant difference in ACC changes in physical reasoning before and after training in the no-ToM training group, with only RT showing a significant decline. There are two possible reasons why ACC did not change in the no-ToM training group before and after training. One reason may be related to the content of the training, which focused on the change of the physical characteristics of the stimuli, such as shape and size, and did not include the outcome of the stimuli following manipulation, such as a hat being blown away by the wind or a ball being bounced away. One reason may be that the difficulty of the physical reasoning included in the test was simple enough for deaf college students to perform well without training. The shortening of RT in both training groups may be due to the fact that participants in both groups were more confident in answering correctly when faced with physical reasoning, resulting in a practice effect, which was demonstrated by steadily increasing speed after several tests.

ToM training has a positive effect on ACC of emotional reasoning, but no effect on RT. ToM training included emotional reasoning training. Participants expressed their emotions and simulated their expressions by imitating the scene portrayed in the story, which helped participants identify representations of various emotional states, thus improving ACC of emotional reasoning. However, this training demonstrated no effect on reducing emotional reasoning RT, possibly because emotional judgment is a form of complex reasoning; complex expression can produce different explanations for causes. Some of the expression pictures in the test material may have been difficult to distinguish in a short time, even after multiple measurements. Therefore, all three tests required participants to spend a certain amount of time to identify the details of the emoticons before they could make judgments, which resulted in no change in RT in the ToM training group among the three tests. Compared with the other three tasks, gender judgment should be the easiest task, and both groups consistently scored higher on gender reasoning in the three tests. This could explain why the two groups did not produce significant differences in both ACC and RT of gender reasoning.

It is found that ToM training has a long term effect on ACC and RT of intentional reasoning, a long term effect on ACC of emotional reasoning, and no long term effect on RT of emotional reasoning. These results suggest that ToM training has a long-term effect on both cognitive ToM and affective ToM. The reason that training has no long-term effect on RT of emotional reasoning is understandable, because ToM training has no effect on emotional reasoning. The reasons for this lack of impact have been discussed earlier.

Limitations and prospects

There are several limitations to this study. The first relates to the participants themselves; participants in this study were all deaf college students from a single school. Future studies can expand the scope of participants, as the level of ToM in the deaf community requires attention. The second factor is related to improving the level of ToM. In this study, interpretation was integrated into the training process as a key point, but when analyzing the results, interpretation was not extracted as a factor for separate analysis to see the magnitude of effect it exerted in the training process. Future research can extract several key factors to improve the level of ToM and construct models to compare the effects of factors. The final limitation relates to the timing of the post-test; since the training was scheduled according to the learning progress of the school, the second post-test can only be scheduled 2 months later, and there was no time to schedule a third post-test to further examine the long-term effects of training. In the future, a third post-test can be administered the next semester. Unfortunately, due to the problem of the number of participants, the hearing status was not analyzed as a single variable in this study, and the relationship between the hearing status and training effect was investigated. We will try to add it into the study in the future.

Summary

To summarize, the addition of ToM training involving mental state conversation can improve ToM level. This research training demonstrated a certain impact on the ability of ToM, including enhancement of intentional and emotional reasoning, and can be implemented in the deaf community to help deaf individuals improve their level of ToM and better adapt to society.

Conclusion

This study designed a training course to improve the level of deaf college students' ToM. The results show that through ToM training, the level of ToM (including cognitive ToM and affective ToM) of deaf college students was significantly improved. Furthermore, the effects of the training were still present 2 months later. This shows that mental state conversation is an important factor restricting ToM level in deaf individuals.

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 Ethics Committee of Tianjin Normal University. 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

YW: Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. XL: Conceptualization, Writing – review & editing. SZ: Data curation, Software, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study is outcome of a research project funded by the National Social Science Foundation of China (grant number 20BYY096) for the period May 2020 to October 2023.

Acknowledgments

We thank Qiang Li for help during the study.

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

Adrian, J., Clemente, R., and Villanueva, L. (2007). Mothers' use of cognitive state verbs in picturebook reading and the development of children's understanding of mind: a longitudinal study. *Child Dev.* 78, 1052–1067. doi: 10.1111/j.1467-8624.2007.01052.x

Ambrose, S. E., Walker, E. A., Unflat-Berry, L. M., Oleson, J. J., and Moeller, M. P. (2015). Quantity and quality of caregivers' linguistic input to 18-month and 3-year-old children who are hard of hearing. *Ear Hear.* 36, 48S–59S. doi: 10.1097/AUD.0000000000000209

Apperly, I. A. (2011). *Mindreaders: The cognitive basis of "theory of mind"*. New York: Psychology Press.

Bailey, P. E., Henry, J. D., and Von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging Ment. Health* 12, 499–503. doi: 10.1080/13607860802224243

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. Psyc.* 42, 241–251. doi: 10.1111/1469-7610.00715

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

Boyes Braem, P., and Sutton-Spence, R. (2001). "Functions of the mouthings in the signing of deaf early and late learners of Swiss German sign language" in *The hands are the head of the mouth: The mouth as articulator in sign language* (Hamburg: Signum), 99–131.

Brothers, L., and Ring, B. A. (1992). Neuroethological framework for the representation of minds. *J. Cognit. Neurosci.* 4, 107–118. doi: 10.1162/jocn.1992.4.2.107

Calderon, R., and Greenberg, M. (2003). *Oxford handbook of deaf studies, language and education Oxford*, UK: Oxford University Press.

Caputi, M., Cugnata, F., and Brombin, C. (2021). Theory of mind and loneliness: effects of a conversation based training at school. *Int. J. Psychol.* 56, 257–265. doi: 10.1002/ijop.12707

Cavallini, E., Ceccato, I., Bertoglio, S., Francescani, A., Vigato, F., Ianes, A. B., et al. (2021). Can theory of mind of healthy older adults living in a nursing home be improved? A randomized controlled trial. *Aging Clin. Exp. Res.* 33, 3029–3037. doi: 10.1007/s40520-021-01811-4

Cavallini, E., Bianco, F., Bottiroli, S., Rosi, A., Vecchi, T., Lecce, S., et al. (2015). Training for generalization in theory of mind: a study with older adults. *Front. Psychol.* 6:1123. doi: 10.3389/fpsyg.2015.01123

Clements, W., Rustin, C. L., and McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: a training study of false belief. *Develop. Sci.* 3, 81–92. doi: 10.1111/1467-7687.00102

De Mulder, H., Hakemulder, F., van den Berghe, R., Klaasen, F., and van Berkum, J. (2017). Effects of exposure to literary narrative fiction: from book smart to street smart. *Sci. Stud. Liter.* 7, 129–169. doi: 10.1075/ssol.7.1.06dem

de Rosnay, M., and Hughes, C. (2006). Conversation and theory of mind: do children talk their way to socio-cognitive understanding? *B. J. Dev. Psychol.* 24, 7–37. doi: 10.1348/026151005X82901

de Villiers, J. G., and de Villiers, P. A. (2014). The role of language in theory of mind development. *Top. Lang. Disord.* 34, 313–328. doi: 10.1097/TLD.0000000000000037

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

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., and Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: individual differences and their antecedents. *Child Dev.* 62, 1352–1366. doi: 10.2307/1130811

Dvash, J., and Shamay-Tsoory, S. G. (2014). Theory of mind and empathy as multidimensional constructs. *Top. Lang. Disord.* 34, 282–295. doi: 10.1097/TLD.0000000000000040

Edwards, L., Marschark, M., Kronenberger, W. G., Crowe, K., and Walton, D. (2021). Inferencing abilities of deaf college students: foundations and implications for metaphor comprehension and theory of mind. *J. Dev. Phys. Disabil.* 33, 233–258. doi: 10.1007/s10882-020-09746-w

Emmorey, K., Thompson, R., and Colvin, R. (2009). Eye gaze during comprehension of American sign language by native and beginning signers. *J. Deaf Stud. Deaf Educ.* 14, 237–243. doi: 10.1093/deafed/enn037

Hao, J., Su, Y., and Chan, R. (2010). Do deaf adults with limited language have advanced theory of mind? *Res. Dev. Disabil.* 31, 1491–1501. doi: 10.1016/j.ridd.2010.06.008

Hynes, C. A., Baird, A. A., and Grafton, S. T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia* 44, 374–383. doi: 10.1016/j.neuropsychologia.2005.06.011

Keil, F. (2006). Explanation and understanding. *Annu. Rev. Psychol.* 57, 227–254. doi: 10.1146/annurev.psych.57.102904.190100

Kidd, D. C., and Castano, E. (2013). Reading literary fiction improves theory of mind. *Science* 342, 377–380. doi: 10.1126/science.1239918

Lecce, S., Caputi, M., and Pagnin, A. (2014). Long-term effect of theory of mind on academic achievement: the role of sensitivity to criticism. *Eur. J. Dev. Psychol.* 11, 305–318. doi: 10.1080/17405629.2013.821944

Lecciso, F., Petrocchi, S., and Marchetti, A. (2013). Hearing mothers and oral deaf children: an atypical relational context for theory of mind. *Eur. J. Psychol. Educ.* 28, 903–922. doi: 10.1007/s10212-012-0146-1

Lecciso, F., Levante, A., Baruffaldi, F., and Petrocchi, S. (2016). Theory of mind in deaf adults. *Cogent Psychol.* 3:1264127. doi: 10.1080/23311908.2016.1264127

Lederberg, A. R., Schick, B., and Spencer, P. E. (2013). Language and literacy development of deaf and hard-of-hearing children: successes and challenges. *Dev. Psychol.* 49, 15–30. doi: 10.1037/a0029558

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

MacDonald, P. M., Kirkpatrick, S. W., and Sullivan, L. A. (1996). Production of facial expressions of emotion in preschool children. *Percept. Mot. Skills* 82, 76–78. doi: 10.2466/pms.1996.82.1.76

Marschark, M., Spencer, P. E., Adams, J., and Sapere, P. (2011). Evidence-based practice in educating deaf and hard-of-hearing children: teaching to their cognitive strengths and needs. *Eur. J. Spec. Needs Edu.* 26, 3–16. doi: 10.1080/08856257.2011.543540

Melot, A. N., and Angeard, N. (2003). Theory of mind: is training contagious? *Developmental Sci.* 6, 178–184. doi: 10.1111/1467-7687.00269

Muir, L. J., and Richardson, I. E. (2002). Video telephony for the deaf: analysis and development of an optimised video compression product. In Proceedings of the Tenth ACM international Conference on Multimedia, 650–652.

Murdaugh, D. L., Nadendla, K. D., and Kana, R. K. (2014). Differential role of temporoparietal junction and medial prefrontal cortex in causal inference in autism: an independent component analysis. *Neurosci. Lett.* 568, 50–55. doi: 10.1016/j.neulet.2014.03.051

O'Reilly, K., Peterson, C. C., and Wellman, H. M. (2014). Sarcasm and advanced theory of mind understanding in children and adults with prelingual deafness. *Dev. Psychol.* 50, 1862–1877. doi: 10.1037/a0036654

Peterson, C. C., and Slaughter, V. (2006). Telling the story of theory of mind: deaf and hearing children's narratives and mental state understanding. *Br. J. Dev. Psychol.* 24, 151–179. doi: 10.1348/026151005X60022

Peterson, C. C., and Slaughter, V. (2003). Opening windows into the mind: mothers' preferences for mental state explanations and children's theory of mind. *Cognit. Dev.* 18, 399–429. doi: 10.1016/S0885-2014(03)00041-8

Peterson, C., and Siegal, M. (1999). Representing inner worlds: theory of mind in autistic, deaf, and normal hearing children. *Psychol. Sci.* 10, 126–129. doi: 10.1111/1467-9280.00119

Petrocchi, S., Filipponi, C., Antonietti, C., Levante, A., and Lecciso, F. (2021). Theory of mind as a mediator between emotional trust beliefs and interpersonal communication competence in a group of young adults. *Psychol. Rep.* 124, 555–576. doi: 10.1177/0033294120913489

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

Rachford, D., and Furth, H. (1986). Understanding friendship and social rules by deaf and hearing children. *J. Appl. Dev. Psychol.* 7, 391–402. doi: 10.1016/0193-3973(86)90008-0

Radecki, M. A., Cox, S. R., and MacPherson, S. E. (2019). Theory of mind and psychosocial characteristics in older men. *Psychol. Aging* 34, 145–151. doi: 10.1037/pag0000324

Schult, C. A., and Wellman, H. M. (1997). Explaining human movements and actions: Children's understanding of the limits of psychological explanation. *Cognition* 62, 291–324. doi: 10.1016/S0010-0277(96)00786-X

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

Shamay-Tsoory, S. G., Tibi-Elhanany, Y., and Aharon-Peretz, J. (2006). The ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Soc. Neurosci.* 1, 149–166. doi: 10.1080/17470910600985589

Slaughter, V., and Peterson, C. C. (2012). *Access to language and cognitive development*. New York: Oxford University Press.

Sutton-Spence, R. (2007). Mouthings and simultaneity in British sign language. *Simul. Sign. Lang.*, 147–162. doi: 10.1075/cilt.281.07sut

Thye, M. D., Ammons, C. J., Murdaugh, D. L., and Kana, R. K. (2018). Differential recruitment of theory of mind brain network across three tasks: an independent component analysis. *Behav. Brain Res.* 347, 385–393. doi: 10.1016/j.bbr.2018.03.041

Tucci, S. L., Easterbrooks, S. R., and Lederberg, A. R. (2016). The effects of theory of mind training on the false belief understanding of deaf and hard-of-hearing students in prekindergarten and kindergarten. *J. Deaf Stud. Deaf Educ.* 21, 310–325. doi: 10.1093/deafed/enw031

Völlm, B. A., Taylor, A. N. W., Richardson, P., Corcoran, R., Stirling, J., McKie, J., et al. (2006). Neuronal correlates theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *NeuroImage* 29, 90–98. doi: 10.1016/j.neuroimage.2005.07.022

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

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

Zhang, H, and Wang, X. (1985). *Raven standard reasoning test manual (Chinese urban revision)*. Beijing: Beijing Normal University Press. 1–60.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 June 2024
DOI 10.3389/fpsyg.2024.1354207

OPEN ACCESS

EDITED BY
Douglas Frye,
University of Pennsylvania, United States

REVIEWED BY
Chi-Lin Yu,
University of Michigan, United States
Carolyn Palmquist,
Amherst College, United States

*CORRESPONDENCE
Honglan Li
✉ honglan.li@njust.edu.cn

RECEIVED 12 December 2023
ACCEPTED 03 June 2024
PUBLISHED 12 June 2024

CITATION
Li H and Leung M-T (2024) The roles of language and executive function in Mandarin-speaking children's theory of mind development.
Front. Psychol. 15:1354207.
doi: 10.3389/fpsyg.2024.1354207

The roles of language and executive function in Mandarin-speaking children's theory of mind development

Honglan Li[1]* and Man-Tak Leung[2]

[1]School of Foreign Studies, Nanjing University of Science of Technology, Nanjing, China, [2]Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University, Kowloon, Hong Kong SAR, China

Introduction: Research has indicated that language and executive function relate closely to first-order false belief reasoning, yet their roles in second-order false belief reasoning are under-explored, and their interplay in theory of mind development remains obscure.

Methods: This study assessed 160 Mandarin-speaking preschoolers' and early primary schoolers' language, executive function, and theory of mind abilities to examine the unique roles and interplay of language and executive function in first-order and second-order false belief reasoning.

Results: Results showed that language significantly uniquely predicted the children's first-order as well as second-order false belief reasoning when controlling for the effects of age and executive function. Although executive function significantly predicted first-order FB reasoning when controlling for age, it was no longer a significant predictor of first-order FB reasoning when language was included in the model. However, executive function played a significant unique role in second-order FB reasoning when controlling for the effects of age and language.

Discussion: The current findings suggest that language plays a greater role than executive function in Mandarin-speaking children's theory of mind development and the contributors to theory of mind development vary in different levels of false belief reasoning.

KEYWORDS

language, executive function, theory of mind, first-order false belief reasoning, second-order false belief reasoning

1 Introduction

Preschoolers experience rapid development in theory of mind (ToM), the ability to understand one's own or others' mental states (e.g., desires, intentions, beliefs and feelings, etc.) and predict others' behaviors according to their mental states (Premack and Woodruff, 1978). Four-year-olds begin to acquire first-order false belief (FB) reasoning (others have different beliefs from their own or reality), the hallmark ability of ToM (Wellman et al., 2001). During 5–6 years old, children begin to understand more advanced second-order FB reasoning (one person's false beliefs about a third person's false beliefs) (Perner and Wimmer, 1985). Regarding the mechanism of ToM development, researchers have proposed that there are multiple routes, including various aspects of language and executive function (EF) for children to develop ToM (Im-Bolter et al., 2016; Farrar et al., 2017). Substantial studies have documented that both language and EF play important roles in ToM development (Milligan et al., 2007; Devine and Hughes, 2014); however, most of previous research focused on the relations between first-order FB reasoning and language or EF, and did not concurrently examine the roles of language and EF in FB reasoning. To date, the effects of language and EF in higher-order ToM are far from clear, and the interplay of language and EF in ToM development remains inconclusive. To better understand how ToM develops as a function of intersections with other cognitive abilities, this study examined the unique roles and interplay of language and EF in Mandarin-speaking children's first-order FB and second-order FB reasoning.

The crucial roles of various aspects of language in ToM development have been well established (Milligan et al., 2007). For instance, general language measured by tasks such as the Reynell Developmental Language Scales (Reynell and Huntley, 1985) and the Test of Early Language Development – 3 (Hresko et al., 1999) that assess children's vocabulary, syntactic and semantic knowledge has been found playing an important role in ToM development (e.g., Cheung et al., 2004; Tardif et al., 2007). More specifically, sentential complement structure has been found playing a crucial role in first-order FB reasoning as well. According to linguistics determinism, sentential complement structure provides a representational format for FB reasoning (de Villiers, 2000). Different from other complex syntactic structures such as relative clauses, sentences with complement clauses, for example, *John thinks that Mary is at the office*, make it possible that the whole sentence is true, while the proposition is false. de Villiers (2000) and de Villiers and Pyers (2002) found that children's earlier sentential complementation predicted their later first-order FB understanding, but not vice versa. Additionally, several training studies have attested to the importance of complementation in first-order FB reasoning (e.g., Lohmann and Tomasello, 2003; Mo et al., 2014).

In recent years, another aspect of language, namely verb factivity, has been found closely related to first-order FB as well as second-order FB reasoning. Verb factivity refers to a feature of predicates that presupposes speakers' beliefs about the truth or falsity of propositions of complement clauses (Kiparsky and Kiparsky, 1971). For instance, *John knew/pretended/thought that the cake was in the box*, the factive verb *know* and the counter-factive verb *pretend* presuppose the speaker's true and false belief about the location of the cake, respectively. By contrast, the non-factive verb *think* does not presuppose the speaker's true or false belief about the location of the cake. Therefore, a comprehensive understanding of verb factivity requires monitoring others' mental states. Specifically, the hearer needs to distinguish his own, the sentence subject's, and the speaker's beliefs from each other. It has been hypothesized that verb factivity conveyed by counter-factive verbs and FB reasoning are naturally related as both involve decoupling a false mental representation of reality (Chen et al., 2012). The findings from neuro-imaging research have indicated that verb factivity and first-order FB reasoning share neural basis (Cheung et al., 2012). Moreover, results from cross-sectional (Cheung et al., 2009; Li and Leung, 2020) as well as longitudinal (Kristen-Antonow et al., 2019; Siu and Cheung, 2022) studies have attested to a close link between verb factivity and first-order FB reasoning.

Apart from first-order FB reasoning, language plays an important role in second-order FB reasoning as well. For example, Hollebrandse et al. (2014) examined whether language supported the development of second-order FB reasoning by comparing 6- to 9-year-old Dutch-speaking children's performance on a verbal and a low-verbal version of a second-order FB task. They found that the children significantly performed better in verbal than in low-verbal second-order FB tasks, which indicated that language might facilitate children's explicit second-order FB understanding. Arslan et al. (2017) attested to a significant correlation between second-order syntactic recursion of relative clauses and second-order FB reasoning after the effects of age and simple working memory were removed. Kristen-Antonow et al. (2019) and Li and Leung (2020) found that verb factivity was closely related to second-order FB reasoning. Although a few studies have explored the relation between language and second-order FB reasoning, they were limited in assessing only one aspect of language, and not taking account of other potential cognitive contributors such as EF. Therefore, a clear picture of the role of language in second-order FB reasoning is yet to emerge.

Executive function (EF) encompasses a set of cognitive processes such as inhibitory control, cognitive flexibility, and working memory that underlie goal-directed behaviors (Diamond, 2013). Inhibitory control involves inhibiting responses to prepotent but irrelevant stimuli while pursuing a cognitively represented goal (Carlson and Moses, 2001). Cognitive flexibility refers to change perspectives based on a shift in rules or demands (Diamond, 2013). Working memory involves both holding and manipulating information in mind (Diamond, 2013). It has been proposed that EF is necessary for the emergence and expression of ToM ability (Carlson and Moses, 2001; Flynn et al., 2004; Benson et al., 2013; Devine and Hughes, 2014). To pass FB tasks, for example, the change-of-location FB tasks, in which participants are required to predict a story protagonist's behavior based on his or her false belief about an object's location (Wimmer and Perner, 1983), participants need to simultaneously hold in mind conflicting perspectives about an object's location, to suppress their own prepotent knowledge of the object's current location, and to predict story protagonist's action according to their less salient false representation of the object's location (Carlson and Moses, 2001). Numerous studies have found robust links between preschoolers' first-order FB understanding and inhibition, cognitive flexibility and working memory independent of age and verbal ability (e.g., Hughes, 1998a; Henning et al., 2011; Devine and Hughes, 2014; Carlson et al., 2015; Duh et al., 2016).

To date, most of previous studies have focused on the relation between EF and first-order FB reasoning in preschoolers, the role of EF in older children's higher-order FB reasoning has been relatively less examined. The literature on the relation between EF and advanced ToM has shown that EF is significantly correlated with advanced ToM in middle childhood (Charman et al., 2001), even when controlling for age and intelligence, or verbal ability (Perner et al., 2002; Devine et al., 2016; Wang et al., 2016). Moreover, findings from several longitudinal studies have indicated that early EF significantly predicts later advanced ToM (e.g., Hughes, 1998b; Austin et al., 2014; Lecce et al., 2017; Shahaeian et al., 2023). However, previous studies have produced inconsistent findings regarding which subcomponents of EF predict advanced ToM. For instance, Hughes (1998b) found that early inhibition significantly predicted later ToM, while Lecce et al. (2017) found that early working memory significantly predicted later ToM, and Austin et al. (2014) found that early attention shifting and working memory but not inhibition significantly predicted later ToM.

The findings from previous studies suggest that the relation between EF and ToM varies in children of different ages and in different levels of FB reasoning. For instance, the correlation between EF and first-order FB reasoning was significant for 3- and 4-year-olds but not for 2-year-olds (Müller et al., 2012). Inhibition rather than working memory significantly predicted 3- to 4-year-olds' first-order FB performance (Carlson et al., 2002); while, working memory was significantly correlated with 4- to 6-year-olds' (Perner et al., 2002) and predicted 4- to 8-year-olds' (Arslan et al., 2017) second-order FB reasoning. Therefore, to better understand the relation between EF and ToM, it is necessary to examine the role of EF in different levels of FB reasoning throughout childhood.

Both language and EF play important roles in ToM development, and these two are closely related as well (Shokrkon and Nicoladis, 2022). For instance, success on the complements task requires one to hold different pieces of information in mind and inhibit a prepotent response to correctly answer a test question, suggesting that working memory and inhibitory control play important roles in language tasks (Tardif et al., 2007). To pass EF tasks such as the Dimensional Change Card Sort (DCCS) task (Zelazo, 2006), a certain level of language capacity is required to understand the *if-if-then* sentence structures. Therefore, EF and language should influence each other's role in ToM development; however, only a few studies have examined the interplay of EF and language in ToM development.

Previous research on the interrelations among language, EF and ToM development have yielded mixed results. A few studies have found that EF fully mediates the role of language in FB reasoning. Low (2010) found that verbal ability was no longer a significant predictor of 3- and 4-year-olds' first-order FB understanding after the effect of cognitive flexibility was removed. Similarly, in Farrant et al.'s (2012) longitudinal study, children's earlier sentential complement ability failed to significantly predict their later FB understanding when controlling for cognitive flexibility. On the contrary, findings from other studies suggest that language may mediate the role of EF in FB reasoning. Jenkins and Astington (1996) found that 3- to 5-year-olds' working memory no longer accounted for any unique variance in their first-order FB reasoning when controlling for language. In a study by Hughes (1998a), the strength of correlations between 3- and 4-year-olds' EF (working memory and inhibition) and first-order FB understanding was no longer significant or reduced, after removing the effects of verbal and non-verbal abilities.

Previous studies on the relations among language, EF and FB reasoning have focused on preschoolers' first-order FB reasoning, the interplay of language and EF in second-order FB reasoning beyond preschoolers is less studied and inconclusive. Moreover, to date, most research focuses on English-speaking children. Although previous studies on the role of language in ToM development in Mandarin-speaking children (e.g., Mo et al., 2014) have yielded similar results to those involving English-speaking children (e.g., Lohmann and Tomasello, 2003), and both Mandarin- and English-speaking children have displayed parallel trajectories in ToM development, the two groups differ in various aspects, such as EF development and the use of think-falsely verbs which are relevant to ToM development (Liu et al., 2008). For instance, Chinese preschoolers demonstrated advanced EF but not advanced ToM when compared with their U.S. counterparts (Sabbagh et al., 2006; Liu et al., 2008), indicating that Chinese preschoolers may not rely heavily on EF to develop their ToM. In Chinese, there are some specific verbs that express false beliefs such as *yǐwéi* 'falsely think', and the use of such words in the test questions of FB tasks has been shown to improve Chinese-speaking children's performance in FB understanding (Lee et al., 1999; Tardif et al., 2004). The daily use of those words or certain specific language structure such as complementation may provide children more opportunities to draw their attention to others' minds, or provide them a scaffolding to represent others' mental states. Children from diverse cultures may vary in the tendency to rely more on different factors in ToM development. To better understand the universality of the mechanisms underlying ToM development as well as specific experiential factors in ToM development, it is necessary to conduct research based on non-Western children speaking non-Indo-European languages and to investigate how language and EF work together in different levels of FB reasoning in children of a wider age range.

This study examined the roles of language and EF in Mandarin-speaking children's first-order and second-order FB understanding. As various aspects of language and EF play crucial roles in ToM development, in this study, we examined three important aspects of language (verbal ability, sentential complement and verb factivity), and EF (cognitive flexibility, inhibitory control and working memory) in preschoolers and early primary schoolers to better capture the roles of language and EF in ToM development. Compared with first-order FB reasoning, the roles of language and EF in second-order FB reasoning are relatively under-explored. Therefore, our first goal was to examine whether the significant effects of language and EF on first-order FB extend to second-order FB reasoning or not. Based on previous meta-analytic studies on the relation between language or EF and first-order FB reasoning (Milligan et al., 2007; Devine and Hughes, 2014), and previous studies on the roles of language and EF in second-order FB or advanced ToM (e.g., Austin et al., 2014; Arslan et al., 2017), we expect to find that language and EF would play significant roles in both first-order FB and second-order FB reasoning. Since Chinese preschoolers demonstrate advanced EF but not advanced ToM when compared with their U.S. counterparts (Sabbagh et al., 2006; Liu et al., 2008), and in Chinese, there are think-falsely verbs which would facilitate ToM development, we also expect to find that language plays a greater role than EF in Mandarin-speaking children's FB reasoning.

In addition, our second goal was to examine whether the roles of language and EF on second-order FB are greater than those on first-order FB reasoning, respectively. Since the stories in second-order FB tasks were more complex, for example, in sentence length and in the number of protagonists, they may be more demanding in language and EF capacity, we expect to find that language and EF play greater roles in second-order FB reasoning than in first-order FB reasoning.

Furthermore, since the interplay of language and EF in first-order FB reasoning remains unclear, and that in second-order FB reasoning is under-explored, our third goal was to examine whether language and EF influence each other's role in first-order and second-order FB reasoning or not. Although previous studies on the interplay among language, EF and FB reasoning have yielded conflicting results, with some studies suggesting that language mediated the role of EF in FB reasoning (e.g., Jenkins and Astington, 1996), while some studies indicating the opposite (e.g., Farrant et al., 2012), we expect to find that language and EF partially mediate each other's role in FB reasoning, since language and EF are closely related, and both have been found playing significant unique roles in ToM development (Milligan et al., 2007; Devine and Hughes, 2014; Arslan et al., 2017; Li and Leung, 2020). The current results will offer additional evidence on the roles and interplay of language and EF in first-order FB, especially in second-order FB reasoning from Mandarin-speaking children.

2 Materials and methods

This study is part of a large project on the language and cognitive development of Mandarin-speaking children. The participants, language and FB tasks in this study were included in previously published research (Li and Leung, 2020). Our current investigation focused on the unique, relative roles and interplay of language and EF in ToM development.

2.1 Participants

In this study, 160 native monolingual Mandarin-speaking children (age range: 50–90 months; mean age = 71 months; *SD* = 11 months, 82 boys) were randomly selected from a public kindergarten and from a public primary school in Shenzhen, a southern city in China. The kindergarten and primary school are in Nanshan and Longhua districts, of which the GDP ranked the first and the fourth out of 11 districts in Shenzhen city in 2022, respectively. We estimated that the social economic status of the children ranged from medium to high. According to teachers' reports, they were free of language and cognitive deficits. Ethical approval for this study was provided by the Human Subjects Ethics Sub-committee at the Hong Kong Polytechnic University, and parent consent forms for the children were obtained before testing. The data from a child was dropped because he missed one first-order FB task. Therefore, subsequent analyses were based on the remaining 159 children (age range: 50–90 months; mean age = 71 months; *SD* = 11 months, 81 boys).

2.2 Materials

2.2.1 Verbal ability measure

The Peabody Picture Vocabulary Test – Revised (PPVT-R) (Sang and Miao, 1990) was used to assess participants' verbal ability. In each test trial, participants were required to select one from four objects or scenes in a picture according to the word they heard. The test was discontinued if the participant failed to correctly answer six trials among eight consecutive trials. There were 175 test trials, and each correct response was scored 1.

2.2.2 Complementation measure

The memory for complements task (de Villiers and Pyers, 2002; Durrleman et al., 2016) was adapted to assess participants' understanding of complementation. Twelve short stories were devised, each described a protagonist who made a mistake or told a lie or had a false belief. Each story was accompanied by two pictures and depicted by three sentences. For example, first, a picture was shown and a test sentence was played on a notebook computer (*Grandma says that there is an egg in the bowl.*), followed by another picture and a sentence (*But look, this is only a ball.*), and then a test question (*What did grandma say was in the bowl?*). Test sentences were constructed with the communication verb *shuō* 'say' as the main clause predicate. Each correct response was scored 1.

2.2.3 Verb factivity measure

Participants' comprehension of verb factivity was assessed by a truth value judgment task (Abbeduto and Rosenberg, 1985). After participants were seated, the experimenter told them that a girl hand puppet, *Xiǎohuā* (placed on the right side of participants) would tell them a short story accompanied by a picture (e.g., *This is Dà Péng. He sees a bottle on the cupboard. Dà Péng does not know that there is apple juice in the bottle.*), and then a teacher hand puppet (placed on the left side of participants) would ask them a test question (e.g., *So, is there any apple juice in the bottle?*). Three buttons, marked by labels of *shì* 'yes', *búshì* 'no' and *kěnéngba* 'maybe', respectively, were placed in front of the participants, and they were required to press one of them to make judgments.

Three verbs examined in the verb factivity test were reported here: *zhīdào* 'know' (a factive verb), *juédé* 'think' (a non-factive verb) and *jiǎzhuāng* 'pretend' (a counter-factive verb). These verbs were used as main clause predicates to construct test sentences with complements in three conditions: (1) affirmative main and complement clause predicates ("+ +" hereafter, e.g., *Mary knew that Paul was at home.*), (2) affirmative main clause predicate and negative complement clause predicate ("+ –" hereafter, e.g., *Mary knew that Paul was not at home.*) and (3) negative main clause predicate and affirmative complement clause predicate ("– +" hereafter, e.g., *Mary did not know that Paul was at home.*). Five test sentences with *zhīdào* 'know' and *juédé* 'think' were constructed in each of the three conditions, and with *jiǎzhuāng* 'pretend' were constructed only in "+ +" and "+ –" conditions, because *jiǎzhuāng* 'pretend' carries a sense of negation itself and is seldom used in negation. Therefore, there were 40 test sentences in total. From the perspective of verb factivity, correct responses to *zhīdào* 'know' in "+ +," "+ –" and "– +" conditions are 'yes', 'no' and 'yes', respectively; to *juédé* 'think' in all three conditions are 'maybe' and to *jiǎzhuāng* 'pretend' in "+ +" and "+ –" conditions are 'no' and 'yes', respectively. Each correct response was scored 1. Test sentences were pseudo-randomized, with the same verb in the same condition occurring no more than two consecutive trials.

2.2.4 Inhibition measure

The day-night stroop task (Gerstadt et al., 1994) was used to assess participants' inhibitory control of the prepotent response of matching a word they say to an object shown. Participants were required to say *day*/*night* for each card showing the moon/the sun (Gerstadt et al., 1994). They received 16 test trials (eight sun cards and eight moon cards) in a fixed pseudo-random order with the same type of card occurring no more than two consecutive trials. No feedbacks were provided during the test. Each correct response was scored 1.

2.2.5 Cognitive flexibility measure

The DCCS task was employed to assess participants' cognitive flexibility, including a standard version (suitable for 2- to 5-year olds) and a border version (suitable for 5- to 7-year-olds) (Zelazo, 2006). Two target cards showing a red boat and a blue rabbit, respectively, were affixed to two bookends in front of the participants. Two sorting trays were placed with an approximately 30 cm-interval in front of the bookends. The standard version required participants to sort six cards (three red rabbit cards and three blue boat cards) according to one dimension (e.g., color) in the pre-switch phase, and then another six according to another dimension (e.g., shape) in the post-switch phase. The order of sorting dimensions in the two phases was counterbalanced. The experimenter repeated the rules before each test trial: "When playing the color/shape game, if the card is red/depicts a boat, then put it here; if the card is blue/depicts a rabbit, then put it here." Participants received the post-switch trials if they passed the pre-switch phase by correctly sorting five or more out of the six pre-switch trials.

Participants received the border version if they passed the standard version by correctly sorting five or more out of the six post-switch trials. The border version included 12 test trials (three red rabbit cards and three blue boat cards with/without a black border), and required participants to sort cards with black border according to color, and those without black border according to shape. The experimenter repeated the rules before each test trial: "If there is a black border, then play the color sorting game; if there is no black border, then play the shape sorting game." Participants had to correctly sort 9 or more out of the 12 trials to pass the border version. Test cards in both versions were in pseudo-random order, with the same card occurring no more than two consecutive trials. During the test, no feedbacks were provided. Participants' performance was scored 0 if they failed the pre-switch phase, 1 if they passed the pre-switch phase, 2 if they passed the post-switch phase, and 3 if they passed the border version (Zelazo, 2006).

2.2.6 Working memory measure

The backward digit span test (Davis and Pratt, 1995) was used to assess participants' working memory. The test included 5 two-digit length practice trials, and 14 two- to eight-digit length test trials, with two trials for each length. Participants received test trials until they correctly responded to one practice trial, the test was discontinued if they failed all five practice trials. For test trials, a correct response to at least one of two digits with the same length led to the next two trials which were longer by one digit. The test was discontinued if participants failed to correctly respond to both trials with the same length. The span (the number of digits in the longest correct sequence, ranging from 0 to 7) was registered.

2.2.7 False belief measures

Participants' first-order FB understanding was assessed by two change-of-location tasks (Wimmer and Perner, 1983) and two unexpected-content tasks (Gopnik and Astington, 1988). In each change-of-location task, there were two protagonists (brother and sister; *Xiǎomíng* and mother). Brother/*Xiǎomíng* put a football/a cake in a box/bowl and then left the scene. Sister/Mother transferred the football/cake into a basket/box when brother/*Xiǎomíng* was absent. After a memory control question (*Where did brother/Xiǎomíng put the football/cake at the very beginning?*) and a reality control question (*Where is the football/cake now?*), participants were asked to predict where brother/son would look for the football/cake when they came back. In each unexpected-content task, participants were asked to guess what contents were in a M&M candy tube/cookie box, and then were shown the real contents after their responses. After a memory control question (*Did you remember what's in the tube/box?*), participants were asked to predict their own and another protagonist's beliefs about the contents of the tube/box before opening the tube/box. Each change-of-location task and unexpected-content task included one and two test questions, respectively. Each correct response was scored 1, and the maximum raw score of first-order FB tasks was 6.

Second-order FB understanding was assessed by four second-order FB tasks: the ice-cream van story (Perner and Wimmer, 1985), the soccer practice story (Miller, 2013a), the cake story (Miller, 2013b) and the hidden toy story (Astington et al., 2002). In each story, Protagonist A (*Chéngcheng*/*Xiǎotāo*/*Zǐxuān*/*Dàwěi*) had a false belief about Protagonist B's (*Fāngfang*/*Hàohao*/*Qíngqing*/*Lìli*) belief about the location of an object or event (ice-cream van/playing football/cake/airplane toy). At the end of each story, participants were asked to predict where Protagonist A (*Chéngcheng*/*Xiǎotāo*/*Zǐxuān*/*Dàwěi*) would look for Protagonist B (*Fāngfang*/*Hàohao*) or the object mentioned (cake/airplane toy) before they correctly responded to all control questions. Each correct response was scored 1, and the maximum raw score for second-order FB tasks was 4.

Traditionally, participants receive a point when they pass both control and test questions in a FB task, and those who fail control questions of a FB task are excluded from data analysis, known as "Exclude System" as described by Sobel and Austerweil (2016) (e.g., Perner et al., 1987), or receive no score even they pass test questions, referred to as "Failure System" by Sobel and Austerweil (2016) (e.g., Astington and Jenkins, 1999). In this study, not all children passed the control questions for all four first-order or second-order FB tasks; some failed the control questions for one, two or three first-order or second-order FB tasks. If we excluded the participants who failed control questions for any one of the FB tasks, the sample size would be significantly reduced. Therefore, we did not exclude the children who failed control questions for FB tasks. If we assigned no scores to participants who failed control questions, we would be unable to distinguish them from those who passed control questions but failed test questions for FB tasks. For example, if child A and child B pass test questions for three out of four second-order FB task, but child A passes control questions for three while child B passes the control questions for four second-order FB tasks. In a traditional "Failure System," both children would receive 3 for second-order FB tasks. Despite receiving the same score, we believe that they differ as least their ability to pass control questions for second-order FB tasks. Therefore, we multiplied the child's raw score on first-order or second-order FB tasks by the proportion of the number of control questions-passed first-order or second-order FB tasks. For child A and child B mentioned above, their final second-order FB score will be calculated as $3*(3/4) = 2.25$ and $3*(4/4) = 3$, respectively. As a result, these two children were distinguishable based on their final scores on second-order FB tasks. Although the FB scores derived from the way used in this study were different from those derived from the conventional "Failure System" way, there were significant correlations between the scores for first-order FB and second-order FB tasks derived from the conventional "Failure System" way and those derived from the current way, respectively [First-order FB tasks: $r(159) = 0.87$, $p < 0.001$; second-order FB tasks: $r(159) = 0.85$, $p < 0.001$].

2.3 Procedure

Participants received the test individually in the respective kindergarten and primary school. The process of the testing was audio-recorded. Preschoolers received all tasks in three 30-min sessions. The first session included verbal ability and complementation tests; the second and third sessions included EF, FB, and factivity tests, with FB and factivity test trials equally distributed in the two sessions. Primary schoolers received all tasks in two 45-min sessions. The first session included verbal ability, complementation, and EF tests; the second session included FB and factivity tasks. First-order FB tasks were always administered before second-order FB tasks in each session. The order of FB and factivity tests was counterbalanced. Complementation and factivity test trials and FB questions could be played three times at the most if participants did not hear them clearly.

3 Results

In this section, we first report descriptive statistics of the data, and then results of simple correlation analyses, and those of path analyses which investigated the direct and indirect effects of language and EF on FB reasoning. The descriptive statistical and correlation analyses were conducted by using SPSS 26.0 software, and path analyses were conducted by using Amos 28.0 software.

3.1 Descriptive statistics

Table 1 shows means and standard deviations of each task. It demonstrates that the children performed at a high level on the cognitive flexibility, inhibition, complements and first-order FB tasks, but not on the verbal ability, verb factivity, working memory and second-order FB tasks. One-way ANOVA revealed that there were no gender differences in the performance on each task. Therefore, gender was not entered in any subsequent analyses.

3.2 Correlations

Table 2 demonstrates correlations among variables tested. It shows that age was significantly correlated with all the other variables, and all subcomponents of language and EF were significantly correlated with first-order FB and second-order FB reasoning. Consistent with previous studies (Shokrkon and Nicoladis, 2022), language was significantly correlated with EF as well. Verbal ability was significantly correlated with cognitive flexibility and working memory, complements was significantly correlated with inhibition and working memory, and verb factivity was significantly correlated with all three subcomponents of EF. Overall, the correlation analyses indicate that language, EF and FB abilities were significantly intercorrelated, providing us preliminary bases to examine the unique roles and interplay of language and EF in first-order FB and second-order FB reasoning.

TABLE 1 Means and standard deviations of each task ($N = 159$).

Variables (range)	Mean	*SD*
Verbal ability (0–175)	92.46	23.25
Complements (0–12)	9.52	3.88
Verb factivity (0–40)	22.85	7.20
Cognitive flexibility (0–3)	2.44	0.56
Inhibition (0–16)	15.14	2.22
Working memory (0–7)	1.87	0.98
First-order FB (0–6)	4.56	1.76
Second-order FB (0–4)	1.32	1.30

3.3 Path analyses

We conducted two path analyses to examine the roles of language and EF in first-order FB and second-order FB reasoning. In each analysis, first-order FB and second-order FB were the dependent variables, language or EF was the independent variable or the mediator according to specific purpose. We performed path analysis by using the direct and indirect effects based on bootstrap procedures (1,000 samples) and bias-corrected bootstrap confidence interval (95%). We computed factor scores for language and EF based on all three aspects of language and EF, respectively, and used them in the analyses. Table 3 presents the factor loadings and communalities for the language and EF items.

Figures 1, 2 illustrate the effects of language and EF on first-order FB and second-order FB reasoning. The results reveal that language played a significant unique role in both first-order and second-order FB reasoning, and the effect of language on first-order FB reasoning ($\beta = 0.42$, $p = 0.002$) was higher than that on second-order FB reasoning ($\beta = 0.33$, $p = 0.001$); whereas, EF played a significant unique role in second-order FB reasoning, but not in first-order FB reasoning. For second-order FB reasoning, language ($\beta = 0.33$, $p = 0.001$) played a greater role than EF ($\beta = 0.24$, $p = 0.002$) in it.

To test whether the effects of language on first-order FB ($\beta = 0.42$, $CI_{95\%} = [0.26;\ 0.59]$) and second-order FB reasoning ($\beta = 0.33$, $CI_{95\%} = [0.16;\ 0.52]$), the effects of EF on first-order FB ($\beta = 0.12$, $CI_{95\%} = [-0.04;\ 0.29]$) and second-order FB reasoning ($\beta = 0.24$, $CI_{95\%} = [0.08;\ 0.39]$), the effects of language ($\beta = 0.42$) and EF ($\beta = 0.12$) on first-order FB reasoning, and the effects of language ($\beta = 0.33$) and EF on second-order FB reasoning ($\beta = 0.24$) were significantly different from each other, we followed the method suggested by Cumming (2009) in which significant differences exist between beta coefficients when the corresponding 95% confidence intervals overlap by less than 50% of the length of one confidence interval arm. For each comparison, we calculated half of the average of the overlapping confidence intervals and added it to the beta weight lower bound estimate. It showed that the appropriate confidence intervals overlapped by more than 50% (see Figure 3), indicating that there were no significant differences between the effects of language on first-order FB and on second-order FB reasoning, or between the effects of EF on first-order FB and on second-order FB reasoning, or between the effects of language and EF on first-order FB reasoning, or between the effects of language and EF on second-order FB reasoning.

We summarized the direct and indirect effects of language and EF in Table 4. It reveals a significant indirect effect of EF on first-order FB reasoning ($\beta = 0.11$, $p = 0.002$), and a significant total effect of EF on first-order FB ($\beta = 0.23$, $p = 0.01$). With the inclusion of the mediator, the effect of EF on first-order FB reasoning was insignificant ($\beta = 0.12$, $p = 0.11$). This shows that language significantly fully mediated the relationship between EF and first-order FB reasoning. In contrast, the indirect effect of language on first-order was insignificant ($\beta = 0.04$, $p = 0.09$), and the total effect of language on first-order FB reasoning was significant ($\beta = 0.42$, $p = 0.002$), and the direct effect of language on first-order FB reasoning was still significant ($\beta = 0.46$, $p = 0.002$) when the mediator was included. This shows that EF did not significantly mediate the role of language in first-order FB reasoning. For second-order FB reasoning, the total, direct and indirect effects of EF and language on it were significant, which indicates that EF and

TABLE 2 Correlations of variables tested (*N* = 159).

Variables	1	2	3	4	5	6	7	8
1. Age	–	0.51**	0.38**	0.52**	0.44**	0.28**	0.38**	0.42**
2. Verbal ability		–	0.30**	0.39**	0.37**	0.15	0.34**	0.38**
3. Complements			–	0.37**	0.16	0.22**	0.31**	0.46**
4. Verb factivity				–	0.34**	0.23**	0.28**	0.38**
5. Cognitive flexibility					–	0.09	0.28**	0.20*
6. Inhibition						–	0.19*	0.20**
7. Working memory							–	0.37**
8. First-order FB								–
9. Second-order FB	0.36**	0.36**	0.31**	0.39**	0.32**	0.17*	0.35**	0.47**

*$p < 0.05$, **$p < 0.01$ (two-tailed).

TABLE 3 Factor loadings and communalities for language and EF items (*N* = 159).

Latent variable	Indicator	Factor loading	Communality
Language	Verb factivity	0.79	0.62
	Verbal ability	0.74	0.55
	Complements	0.73	0.53
EF	Working memory	0.78	0.49
	Cognitive flexibility	0.70	0.29
	Inhibition	0.54	0.61

TABLE 4 Summary of the results of path analyses (*N* = 159).

Effect Path	Standardized effect coefficients		
	Total	Direct	Indirect
EF → Language → First-order FB	0.23*	0.12	0.11**
Language → EF → First-order FB	0.46**	0.42**	0.04
EF → Language → Second-order FB	0.33**	0.24**	0.05*
Language → EF → Second-order FB	0.41**	0.33**	0.09**

*$p < 0.05$, **$p < 0.01$ (two-tailed).

language significantly partially mediated each other's role in second-order FB reasoning.

4 Discussion

The current findings partially support our hypotheses. We found that language significantly uniquely predicted Mandarin-speaking children's first-order and second-order FB reasoning when controlling for the effects of age and EF. EF played a significant unique role in second-order FB reasoning, but not in first-order FB reasoning. Regarding the interplay of language and EF in FB reasoning, language significantly fully mediated the effect of EF on first-order FB reasoning, and language and EF significantly partially mediated each other's role in second-order FB reasoning.

4.1 Language and false belief reasoning

Our results are consistent with previous studies in that language played a significant unique role in first-order FB reasoning (e.g., de Villiers and Pyers, 2002; Cheung et al., 2009; Li and Leung, 2020), even when controlling for the effect of EF (e.g., Tardif et al., 2007). In addition, the current finding offers more evidence to the facilitating role of language in second-order FB reasoning (e.g., Hollebrandse et al., 2014; Arslan et al., 2017). Since the stories in second-order FB tasks were more complex than those in first-order FB tasks, we initially anticipated that language would play a greater role in second-order FB reasoning than in first-order FB reasoning. However, contrary to our expectation, we found that language played a greater role in first-order FB reasoning than in second-order FB reasoning. This finding suggests that the children in the current sample leaned more heavily on language to understand first-order FB reasoning than to understand second-order FB reasoning. The crucial role of language in first-order FB development has been well established in existing literature. Longitudinal and training studies have indicated that language plays a pivotal role in aiding children to build representation of other's mental states (e.g., de Villiers and Pyers, 2002; Lockl and Schneider, 2007; Mo et al., 2014). Compared with first-order FB reasoning, second-order FB tasks are more complex in stories involved, and the primary distinction between first-order FB reasoning and second-order FB reasoning might lie in the information-processing requirements of the tasks. Lockl and Schneider (2007) proposed that once children have acquired basic language abilities that help them to develop first-order FB reasoning, their performance on second-order belief reasoning might largely depend on their comprehension of the complex stories used in second-order FB tasks. According to Lockl and Schneider's (2007) proposal, the nature of the role of language in first-order FB reasoning may differ from that in second-order FB reasoning. Specifically, language supports the emergence of first-order FB reasoning, and the expression of second-order FB reasoning (Lockl and Schneider, 2007; Polyanskaya et al., 2022). Nevertheless, to date, the role of language in the transition from first-order FB to

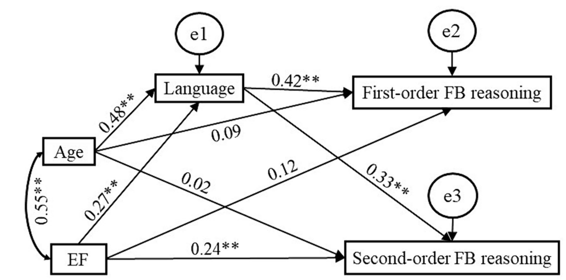


FIGURE 1
Path diagram for the indirect effect of language on the EF-first-order FB reasoning and EF-second-order FB relations. $*p < 0.05$, $**p < 0.01$.

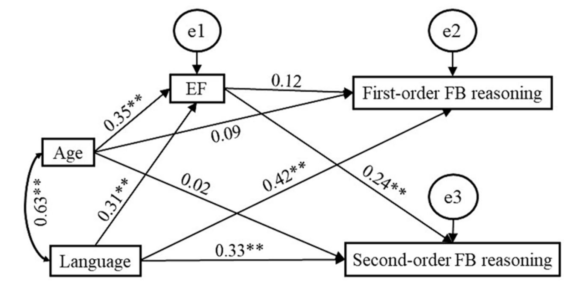


FIGURE 2
Path diagram for the indirect effect of EF on the language-first-order FB reasoning and language-second-order FB relations. $*p < 0.05$, $**p < 0.01$.

second-order FB reasoning remains unclear, and further research is needed to investigate this issue in more detail.

4.2 Executive function and false belief reasoning

Out of our expectation, EF did not play a significant unique role in first-order FB reasoning in this study. This finding is inconsistent with those from previous studies which found a significant predicting role of EF in first-order FB reasoning (e.g., Henning et al., 2011; Duh et al., 2016). The use of different controlled variables in path analyses may be one of the explanations for the discrepancies. In this study, we assessed three important aspects of language: verbal ability, syntactic complements and semantic verb factivity, and all three aspects of language were controlled in the path analyses predicting the role of EF in FB reasoning. By contrast, in previous studies, the controlled variables only included one aspect of language such as verbal ability (Carlson et al., 2015; Lecce et al., 2017) or syntactic competence (Henning et al., 2011), or no language competence (Duh et al., 2016). When controlling for age only, but not for language, we also found that EF significantly predicted first-order FB reasoning, which is in line with previous studies (e.g., Austin et al., 2014; Duh et al., 2016).

Another possible explanation for the discrepancies may be attributed to the variations in the age of the children across this and previous studies. The children in this study were older (mean age: 71 months) than those in previous studies (mean age: 48 to 59 months) (e.g., Henning et al., 2011; Carlson et al., 2015). It is noteworthy that we did not included children younger than 4 years old. Given that findings from previous studies have demonstrated diverse associations

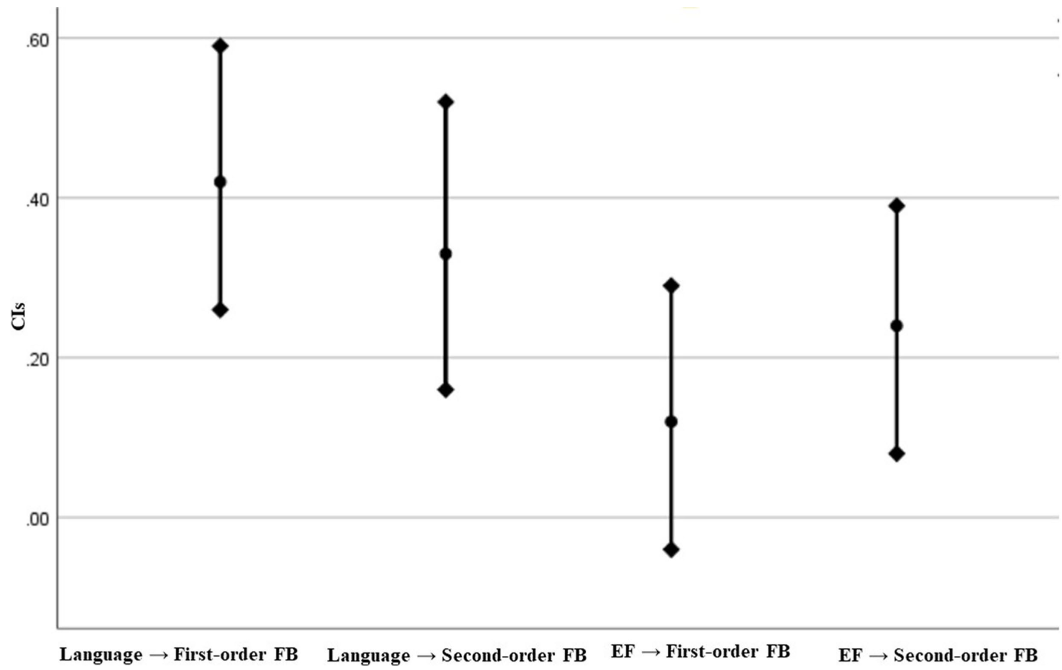


FIGURE 3
Difference between standardized beta coefficients.

between EF and first-order FB reasoning among children of different age groups (Müller et al., 2012), the inclusion of younger children comparable to those in previous studies may yield a different picture of the relation between EF and first-order FB reasoning.

Although EF did not play a significant unique role in Mandarin-speaking children's first-order, it significantly predicted their second-order FB reasoning after the effects of age and language were removed. Our findings indicate that EF played a greater effect on Mandarin-speaking children's second-order FB reasoning than on their first-order FB reasoning, which lends support to Devine and Hughes's (2014) proposal that the correlation strength between EF and FB tasks is contingent upon on the EF load required by the tasks. Specially, FB tasks demanding higher levels of EF are anticipated to exhibit stronger correlation with EF. In this study, second-order FB stories were more complex than first-order FB stories in the number and length of sentences, and in the number of protagonists, thus requiring participants to take more effect to retain and manipulate the story details, to switch flexibly among different protagonists' mental states, and to suppress their own or others' false beliefs. Therefore, the current second-order FB tasks impose more EF demands than first-order FB tasks, leading to an expectedly stronger correlation with EF.

4.3 Language, executive function and false belief reasoning

Regarding the relations among language, EF and FB reasoning, we did not find a mediation role of EF on the relation between language and first-order FB reasoning as previous studies (Low, 2010; Farrant et al., 2012). Instead, the role of EF in first-order FB reasoning was completely mediated by language in this study, which is in line with the findings from Jenkins and Astington (1996) and Hughes (1998b). However, for second-order FB reasoning, as expected, language and EF partially mediated each other's role in it, and the effect of language was greater than that of EF. Together, we found that language played a greater role than EF in 4- to 7-year-old Mandarin-speaking children's first-order and second-order FB reasoning.

The current findings that language played a greater role than EF in Mandarin-speaking children's first-order and second-order FB reasoning support our hypothesis. Chinese children exhibit higher levels of EF than their U.S. counterparts (Sabbagh et al., 2006); however, their ToM is comparable to their U.S. counterparts' (Liu et al., 2008). The findings of Sabbagh et al. (2006) and Liu et al. (2008), together suggest that Chinese children may not heavily rely on EF to develop their ToM. The current findings indicate that Chinese children relied more on language than on EF in FB reasoning. A possible explanation may be that in Chinese, there are some specific verbs that express false beliefs such as *yǐwéi* 'falsely think,' and the use of those words in everyday conversations directly exposes children to instances of false beliefs, enabling them to observe others' mental states and understand diverse perspectives on an event. This exposure would play a facilitating role in developing their ToM ability. However, the underlying reasons for why Chinese children tend to rely more on language to develop ToM warrant further cross-cultural research.

While this study focused on typically developing children, our findings align with research involving deaf and hard-hearing (DHH) children in that language plays an important role in FB reasoning. The literature on DHH children has shown that language is a key factor in their ToM development (e.g., Schick et al., 2007). Specifically, DHH children born to hearing parents often experience delayed language exposure, which correlates with their delays in ToM development (e.g., Stanzione and Schick, 2014; Walker et al., 2017). In contrast, DHH children who receive cochlear implants early or have access to signed language from a young age, have greater access to language and conversational experiences (Dettman et al., 2007), and develop ToM comparable to their hearing peers (e.g., Schick et al., 2007; Sundqvist et al., 2014; Yu et al., 2021). Beyond ToM, language also plays an important role in the EF development of DHH children (e.g., Jones et al., 2020; Goodwin et al., 2022). Studies have demonstrated that language mediates the EF differences between DHH and hearing children (e.g., Botting et al., 2017; Merchán et al., 2022). Despite this, few studies, to our best knowledge to date, have thoroughly explored the interplay between language and EF in the ToM development in DHH children. Therefore, the extent to which language influences the role of EF in ToM development in DHH children or the reverse remains an open question.

Our findings suggest that language plays a greater role than EF in FB reasoning in monolingual children, but the scenario may be different for bilingual children. Previous studies have shown that bilingual children often score lower on language tests than their monolingual counterparts (e.g., Carlson and Meltzoff, 2008; Bialystok and Viswanathan, 2009; Diaz and Farrar, 2018). However, they demonstrate an advantage in ToM development over their monolingual peers (e.g., Goetz, 2003; Farhadian et al., 2010; Schroeder, 2018), and this bilingual advantage in ToM may be explained by a bilingual advantage in EF (e.g., Bialystok and Senman, 2004; Kovács, 2009; Buac and Kaushanskaya, 2020). Therefore, the relative roles of language and EF in FB reasoning may vary between bilingual and monolingual children.

4.4 Limitations

In spite of the contributions to the literature on the relations among language, EF and ToM development, this study has several limitations. The first is that the day-night stroop task is too simple for the children in this study, as their performance on this task approached to the ceiling. In further research, more appropriate tasks should be employed to assess participants' EF ability. The second limitation is that the use of verbal EF and FB tasks. Although we controlled for the effect of language when examining the role of EF in FB reasoning, the use of low-verbal or non-verbal EF and FB tasks would enable us to better elucidate the relation between EF and FB reasoning. The third is that the participants' age range is limited (from ages 4 to 7), thus the current conclusions may not generalize to the roles of EF and language in ToM development before or beyond this age range. In this study, even the older children performed poorly on verb factivity, working memory and second-order FB tasks. The inclusion of children with a wider age range may contribute to a better investigation on the roles of language and EF in higher-order FB reasoning in further research. In addition, social environmental factors such as the number of siblings, economic social status and parents' education background that may contribute to the relations among language, EF and ToM development were not available in this study. The current sample was collected from a kindergarten and a primary school in Shenzhen city. Due to the limited scope of this sample, its representativeness may be insufficient, thereby limiting the generalizability of the current findings to other samples characterized by distinct social environmental factors.

In sum, this study added to our understanding of the roles of language and EF and how they work together in ToM development. The current findings suggest that language plays a more prominent role than EF in 4- to 7-year-old Mandarin-speaking children's first-order and second-order FB reasoning. In this study, we only examined the roles of language and EF in ToM development. In the course of children's language, EF and ToM development, ToM could also play a predictive role in language and EF development. For further research, training or longitudinal studies would allow stronger inferences to the roles and interactions of language and EF in ToM development, and the directionality of effects among language, EF and ToM.

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 the Human Subjects Ethics Sub-committee at the Hong Kong Polytechnic University (reference number: HSEARS20161103002). 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. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

HL: Writing – original draft, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. M-TL: Writing – review & editing, Supervision, Conceptualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Fundamental Research Funds for the Central Universities (No.30921012206), the Research Funds of Philosophy and Social Sciences of the Higher Education Institutions of Jiangsu Province (2021SJA0014), and the Research Funds of the Development of Disability Services of Jiangsu Province (2021SC03005).

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

References

Abbeduto, L., and Rosenberg, S. (1985). Children's knowledge of the presuppositions of know and other cognitive verbs. *J. Child Lang.* 12, 621–641. doi: 10.1017/S0305000900006693

Arslan, B., Hohenberger, A., and Verbrugge, R. (2017). Syntactic recursion facilitates and working memory predicts recursive theory of mind. *PLoS One* 12:e0169510. doi: 10.1371/journal.pone.0169510

Astington, J. W., and Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Dev. Psychol.* 35, 1311–1320. doi: 10.1037/0012-1649.35.5.1311

Astington, J. W., Pelletier, J., and Homer, B. (2002). Theory of mind and epistemological development: the relation between children's second-order false-belief understanding and their ability to reason about evidence. *New Ideas Psychol.* 20, 131–144. doi: 10.1016/S0732-118X(02)00005-3

Austin, G., Groppe, K., and Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a one-year longitudinal perspective. *Front. Psychol.* 5:655. doi: 10.3389/fpsyg.2014.00655

Benson, J. E., Sabbagh, M. A., Carlson, S. M., and Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Dev. Psychol.* 49, 1615–1627. doi: 10.1037/a0031056

Bialystok, E., and Senman, L. (2004). Executive processes in appearance–reality tasks: the role of inhibition of attention and symbolic representation. *Child Dev.* 75, 562–579. doi: 10.1111/j.1467-8624.2004.00693.x

Bialystok, E., and Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112, 494–500. doi: 10.1016/j.cognition.2009.06.014

Botting, N., Jones, A., Marshall, C., Denmark, T., Atkinson, J., and Morgan, G. (2017). Non-verbal executive function is mediated by language: a study of deaf and hearing children. *Child Dev.* 88, 1689–1700. doi: 10.1111/cdev.12659

Buac, M., and Kaushanskaya, M. (2020). Predictors of theory of mind performance in bilingual and monolingual children. *Int. J. Biling.* 24, 339–359. doi: 10.1177/1367006919826866

Carlson, S. M., Claxton, L. J., and Moses, L. J. (2015). The relation between executive function and theory of mind is more than skin deep. *J. Cogn. Dev.* 16, 186–197. doi: 10.1080/15248372.2013.824883

Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298. doi: 10.1111/j.1467-7687.2008.00675.x

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

Carlson, S. M., Moses, L. J., and Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant Child Dev.* 11, 73–92. doi: 10.1002/icd.298

Charman, T., Carroll, F., and Sturge, C. (2001). Theory of mind, executive function and social competence in boys with ADHD. *Emot. Behav. Difficult.* 6, 31–49. doi: 10.1080/13632750100507654

Chen, L., Cheung, H., Szeto, C.-Y., Zhu, Z., and Wang, S. (2012). Do false belief and verb non-factivity share similar neural circuits? *Neurosci. Lett.* 510, 38–42. doi: 10.1016/j.neulet.2011.12.065

Cheung, H., Chen, L., Szeto, C.-Y., Feng, G., Lu, G., Zhang, Z., et al. (2012). False belief and verb non-factivity: a common neural basis? *Int. J. Psychophysiol.* 83, 357–364. doi: 10.1016/j.ijpsycho.2011.12.002

Cheung, H., Chen, H.-C., and Yeung, W. (2009). Relations between mental verb and false belief understanding in Cantonese-speaking children. *J. Exp. Child Psychol.* 104, 141–155. doi: 10.1016/j.jecp.2009.05.004

Cheung, H., Hsuan-Chih, C., Creed, N., Ng, L., Wang, S., and Mo, L. (2004). Relative roles of general and complementation language in theory-of-mind development:

evidence from Cantonese and English. *Child Dev.* 75, 1155–1170. doi: 10.1111/j.1467-8624.2004.00731.x

Cumming, G. (2009). Inference by eye: Reading the overlap of independent confidence intervals. *Stat. Med.* 28, 205–220. doi: 10.1002/sim.3471

Davis, H. L., and Pratt, C. (1995). The development of children's theory of mind: the working memory explanation. *Aust. J. Psychol.* 47, 25–31. doi: 10.1080/00049539508258765

de Villiers, J. G. (2000). "Language and theory of mind: what are the developmental relationships?" in *Understanding other minds: perspectives from developmental cognitive neuroscience*. eds. S. Baron-Cohen, H. Tager-Flusberg and D. J. Cohen. *2nd* ed (Oxford: Oxford University Press), 83–123.

de Villiers, J. G., and Pyers, J. E. (2002). Complements to cognition: a longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cogn. Dev.* 17, 1037–1060. doi: 10.1016/S0885-2014(02)00073-4

Dettman, S. J., Pinder, D., Briggs, R. J. S., Dowell, R. C., and Leigh, J. R. (2007). Communication development in children who receive the cochlear implant younger than 12 months: risks versus benefits. *Ear Hear.* 28, 11S–18S. doi: 10.1097/AUD.0b013e31803153f8

Devine, R. T., and Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: a meta-analysis. *Child Dev.* 85, 1777–1794. doi: 10.1111/cdev.12237

Devine, R., 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, 758–771. doi: 10.1037/dev0000105

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

Diaz, V., and Farrar, M. J. (2018). The missing explanation of the false-belief advantage in bilingual children: a longitudinal study. *Dev. Sci.* 21:e12594. doi: 10.1111/desc.12594

Duh, S., Paik, J. H., Miller, P. H., Gluck, S. C., Li, H., and Himelfarb, I. (2016). Theory of mind and executive function in Chinese preschool children. *Dev. Psychol.* 52, 582–591. doi: 10.1037/a0040068

Durrleman, S., Burnel, M., Thommen, E., Foudon, N., Sonié, S., Reboul, A., et al. (2016). The language cognition interface in ASD: complement sentences and false belief reasoning. *Res. Autism Spectr. Disord.* 21, 109–120. doi: 10.1016/j.rasd.2015.10.003

Farhadian, M., Abdullah, R., Mansor, M., Redzuan, M., Gazanizadand, N., and Kumar, V. (2010). Theory of mind in bilingual and monolingual preschool children. *J. Psychol.* 1, 39–46. doi: 10.1080/09764224.2010.11885444

Farrant, B. M., Maybery, M. T., and Fletcher, J. (2012). Language, cognitive flexibility, and explicit false belief understanding: longitudinal analysis in typical development and specific language impairment. *Child Dev.* 83, 223–235. doi: 10.1111/j.1467-8624.2011.01681.x

Farrar, M. J., Seung, H. K., and Lee, H. (2017). Language and false-belief task performance in children with autism spectrum disorder. *J. Speech Lang. Hear. Res.* 60, 1999–2013. doi: 10.1044/2017_JSLHR-L-15-0422

Flynn, E., O'Malley, C., and Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Dev. Sci.* 7, 103–115. doi: 10.1111/j.1467-7687.2004.00326.x

Gerstadt, C. L., Hong, Y. J., and Diamond, A. (1994). The relationship between cognition and action: performance of children 3½–7 years old on a stroop-like day-night test. *Cognition* 53, 129–153. doi: 10.1016/0010-0277(94)90068-X

Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Biling. Lang. Congn.* 6, 1–15. doi: 10.1017/S1366728903001007

Goodwin, C., Carrigan, E., Walker, K., and Coppola, M. (2022). Language not auditory experience is related to parent-reported executive functioning in preschool-aged deaf and hard-of-hearing children. *Child Dev.* 93, 209–224. doi: 10.1111/cdev.13677

Gopnik, A., and Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Dev.* 59, 26–37. doi: 10.2307/1130386

Henning, A., Spinath, F. M., and Aschersleben, G. (2011). The link between preschoolers' executive function and theory of mind and the role of epistemic states. *J. Exp. Child Psychol.* 108, 513–531. doi: 10.1016/j.jecp.2010.10.006

Hollebrandse, B., van Hout, A., and Hendriks, P. (2014). Children's first and second-order false-belief reasoning in a verbal and a low-verbal task. *Synthese* 191, 321–333. doi: 10.1007/s11229-012-0169-9

Hresko, W. P., Reid, D. K., and Hammill, D. S. (1999). *The test of early language development 3 (TELD-3)*. Austin, TX: Pro-Ed.

Hughes, C. (1998a). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Dev. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Hughes, C. (1998b). Finding your marbles: does preschoolers' strategic behavior predict later understanding of mind? *Dev. Psychol.* 34, 1326–1339. doi: 10.1037/0012-1649.34.6.1326

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

Jenkins, J. M., and Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Dev. Psychol.* 32, 70–78. doi: 10.1037/0012-1649.32.1.70

Jones, A., Atkinson, J., Marshall, C., Botting, N., St Clair, M. C., and Morgan, G. (2020). Expressive vocabulary predicts non-verbal executive function: a 2-year longitudinal study of deaf and hearing children. *Child Dev.* 91, e400–e414. doi: 10.1111/cdev.13226

Kiparsky, P., and Kiparsky, C. (1971). "Fact" in *Semantics*. eds. D. D. Steinberg and L. A. Jakobovits (New York: Cambridge University Press), 345–369.

Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Dev. Sci.* 12, 48–54. doi: 10.1111/j.1467-7687.2008.00742.x

Kristen-Antonow, S., Jarvers, I., and Sodian, B. (2019). Preschoolers' developing understanding of factivity in mental verb comprehension and its relation to first- and second-order false belief understanding: a longitudinal study. *J. Cogn. Dev.* 20, 354–369. doi: 10.1080/15248372.2019.1586710

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

Lee, K., Olson, D. R., and Torrance, N. (1999). Chinese children's understanding of false beliefs: the role of language. *J. Child Lang.* 26, 1–21. doi: 10.1017/S0305000998003626

Li, H., and Leung, M.-T. (2020). Relations between verb factivity and first-order and second-order false belief understanding: evidence from mandarin-speaking typically developing children and children with autism spectrum disorders. *Clin. Linguist. Phon.* 34, 185–200. doi: 10.1080/02699206.2019.1628810

Liu, D., Wellman, H. M., Tardif, T., and Sabbagh, M. A. (2008). Theory of mind development in Chinese children: a meta-analysis of false-belief understanding across cultures and languages. *Dev. Psychol.* 44, 523–531. doi: 10.1037/0012-1649.44.2.523

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

Lohmann, H., and Tomasello, M. (2003). The role of language in the development of false belief understanding: a training study. *Child Dev.* 74, 1130–1144. doi: 10.1111/1467-8624.00597

Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: relations with complex syntactical mastery. *Child Dev.* 81, 597–615. doi: 10.1111/j.1467-8624.2009.01418.x

Merchán, A., Fernández García, L., Gioiosa Maurno, N., Ruiz Castañeda, P., and Daza González, M. T. (2022). Executive functions in deaf and hearing children: the mediating role of language skills in inhibitory control. *J. Exp. Child Psychol.* 218:105374. doi: 10.1016/j.jecp.2022.105374

Miller, S. A. (2013a). Children's understanding of second-order false belief: comparisons of content and method of assessment. *Infant Child Dev.* 22, 649–658. doi: 10.1002/icd.1810

Miller, S. A. (2013b). Effects of deception on children's understanding of second-order false belief. *Infant Child Dev.* 22, 422–429. doi: 10.1002/icd.1799

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

Mo, S., Su, Y., Sabbagh, M. A., and Xiu, J. (2014). Sentential complements and false belief understanding in Chinese mandarin-speaking preschoolers: a training study. *Cogn. Dev.* 29, 50–61. doi: 10.1016/j.cogdev.2013.11.001

Müller, U., Liebermann-Finestone, D. P., Carpendale, J. I., Hammond, S. I., and Bibok, M. B. (2012). Knowing minds, controlling actions: the developmental relations between theory of mind and executive function from 2 to 4 years of age. *J. Exp. Child Psychol.* 111, 331–348. doi: 10.1016/j.jecp.2011.08.014

Perner, J., Kain, W., and Barchfeld, P. (2002). Executive control and higher-order theory of mind in children at risk of ADHD. *Infant Child Dev.* 11, 141–158. doi: 10.1002/icd.302

Perner, J., Leekam, S. R., and Wimmer, H. (1987). Three-year-olds' difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5, 125–137. doi: 10.1111/j.2044-835X.1987.tb01048.x

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

Polyanskaya, I., Eigsti, I.-M., Brauner, T., and Blackburn, P. (2022). Second-order false beliefs and linguistic recursion in autism spectrum disorder. *J. Autism Dev. Disord.* 52, 3991–4006. doi: 10.1007/s10803-021-05277-1

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

Reynell, J., and Huntley, M. (1985). *Reynell developmental language scales*. *2nd* Edn. Windsor, England: NFER-Nelson.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., and Lee, K. (2006). The development of executive functioning and theory of mind: a comparison of Chinese and U.S. preschoolers. *Psychol. Sci.* 17, 74–81. doi: 10.1111/j.1467-9280.2005.01667.x

Sang, B., and Miao, X. (1990). Pi Bodi tupian cihui ceyan xiudingban (PPVT-R) Shanghai shiqu shiyong changmo de xiuding [the revision of trail norm of Peabody picture vocabulary test revised (PPVT-R) in Shanghai proper]. *J. Psychol. Sci.* 14, 20–25.

Schick, B., De Villiers, P., De Villiers, J., and Hoffmeister, R. (2007). Language and theory of mind: a study of deaf children. *Child Dev.* 78, 376–396. doi: 10.1111/j.1467-8624.2007.01004.x

Schroeder, S. R. (2018). Do bilinguals have an advantage in theory of mind? A meta-analysis. *Front. Communicat.* 3:36. doi: 10.3389/fcomm.2018.00036

Shahaeian, A., Haynes, M., and Frick, P. J. (2023). The role of language in the association between theory of mind and executive functioning in early childhood: new longitudinal evidence. *Early Child. Res. Q.* 62, 251–258. doi: 10.1016/j.ecresq.2022.09.003

Shokrkon, A., and Nicoladis, E. (2022). The directionality of the relationship between executive functions and language skills: a literature review. *Front. Psychol.* 13:848696. doi: 10.3389/fpsyg.2022.848696

Siu, C. T.-S., and Cheung, H. (2022). A longitudinal reciprocal relation between theory of mind and language. *Cogn. Dev.* 62:101176. doi: 10.1016/j.cogdev.2022.101176

Sobel, D. M., and Austerweil, J. L. (2016). Coding choices affect the analyses of a false belief measure. *Cogn. Dev.* 40, 9–23. doi: 10.1016/j.cogdev.2016.08.002

Stanzione, C., and Schick, B. (2014). Environmental language factors in theory of mind development: evidence from children who are deaf/hard-of-hearing or who have specific language impairment. *Top. Lang. Disord.* 34, 296–312. doi: 10.1097/TLD.0000000000000038

Sundqvist, A., Lyxell, B., Jönsson, R., and Heimann, M. (2014). Understanding minds: early cochlear implantation and the development of theory of mind in children with profound hearing impairment. *Int. J. Pediatr. Otorhinolaryngol.* 78, 538–544. doi: 10.1016/j.ijporl.2013.12.039

Tardif, T., So, C. W.-C., and Kaciroti, N. (2007). Language and false belief: evidence for general, not specific, effects in Cantonese-speaking preschoolers. *Dev. Psychol.* 43, 318–340. doi: 10.1037/0012-1649.43.2.318

Tardif, T., Wellman, H. M., and Cheung, K. M. (2004). False belief understanding in Cantonese-speaking children. *J. Child Lang.* 31, 779–800. doi: 10.1017/S0305000904006580

Walker, E. A., Ambrose, S. E., Oleson, J., and Moeller, M. P. (2017). False belief development in children who are hard of hearing compared with peers with normal hearing. *J. Speech Lang. Hear. Res.* 60, 3487–3506. doi: 10.1044/2017_JSLHR-L-17-0121

Wang, Z., Devine, R. T., Wong, K. K., and Hughes, C. (2016). Theory of mind and executive function during middle childhood across cultures. *J. Exp. Child Psychol.* 149, 6–22. doi: 10.1016/j.jecp.2015.09.028

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

Yu, C.-L., Stanzione, C. M., Wellman, H. M., and Lederberg, A. R. (2021). Theory-of-mind development in young deaf children with early hearing provisions. *Psychol. Sci.* 32, 109–119. doi: 10.1177/0956797620960389

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

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 June 2024
DOI 10.3389/fpsyg.2024.1348074

OPEN ACCESS

EDITED BY
Douglas Frye,
University of Pennsylvania, United States

REVIEWED BY
James Hugo Smith-Spark,
London South Bank University,
United Kingdom
Munis Dundar,
Erciyes University, Türkiye

*CORRESPONDENCE
Leyre Gambra
✉ lgambra@unav.es
Nerea-Crespo-Eguílaz
✉ nerea.crespo@unir.net

RECEIVED 04 December 2023
ACCEPTED 23 May 2024
PUBLISHED 12 June 2024

CITATION
Gambra L, Magallon S and Crespo-Eguílaz N (2024) Weak central coherence in neurodevelopmental disorders: a comparative study. *Front. Psychol.* 15:1348074. doi: 10.3389/fpsyg.2024.1348074

Weak central coherence in neurodevelopmental disorders: a comparative study

Leyre Gambra[1]*, Sara Magallon[1] and Nerea Crespo-Eguílaz[2,3]*

[1]Faculty of Education and Psychology, University of Navarra, Pamplona, Spain, [2]Pediatric Neurology Unit, Clinic University of Navarra, Pamplona, Spain, [3]Facultad de Educación, Universidad Internacional de La Rioja, Logroño, Spain

Introduction: Central coherence is the normal tendency to process and give meaning to incoming information taking into account the context or global view of that information.

Methods: We assessed the central coherence of 252 school children of normal intelligence between 6 and 11 years old. We compared the performance of two groups: (a) a control group (n = 194), and (b) a clinical group (n = 58) comprising children with NVLD+ADHD (n = 24), ADHD alone (n = 16), SCD (n = 8) and level-1ASD (n = 10) (Kluskall-Wallis H and Mann-Whitney U were calculated to make comparisons within groups and between pairs of groups). The effects of medication were studied (Student's t test).

Results: The NVLD+ADHD, SCD and ASD1 groups showed weak central coherence. The performance of the ADHD group was normal and differed significantly from the NVLD+ADHD group.

Conclusion: Central coherence deficit was not exclusive to ASD1: it also characterizes NVLD and SCD.

KEYWORDS

central coherence, non-verbal learning disorder, attention deficit and hyperactivity disorder, social communication disorder, autism

1 Introduction

The world is perceived as being hierarchically organized and includes global perceptions comprising local details (D'Souza et al., 2016); a human being is able to process information at both global and local levels. This capacity, which is denominated central coherence, is implied in everyday activities such as categorization, inspection of details in our surroundings, perception of the structure of a scene and analysis of information (Nayar et al., 2017). Central coherence is, therefore, the normal tendency to process and give meaning, in a global manner, to incoming information in its context (Noens and van Berckelaer-Onnes, 2008). It is thanks to central coherence that when we receive a message, we prioritize comprehension of meaning and not just the literal form (Crespo-Eguílaz et al., 2012).

Attwood (Attwood, 2007) provides a metaphor that may be useful to understand the nature of weak central coherence: "imagine rolling up a sheet of paper to form a tube and with one eye closed bring it up to other open eye, as if it were a telescope, and look at the world through it: you see the details but do not perceive the context." A person with difficulty in central coherence (that is, a preference for analytical processing rather than global processing) has specific difficulties in simultaneously processing information perceived and in giving coherent, integral meaning to that information. The cognitive style of such a person is, therefore, characterized by a tendency to process details (Navon, 1977; Schooler, 2002; Förster and Dannenberg, 2010; Crespo-Eguílaz and Narbona, 2011) This style of processing is slower and more demanding from the cognitive point of view (Nayar et al., 2017). An ability to carry out tasks requiring central coherence rapidly is of fundamental importance in learning and in social behavior (Crespo-Eguílaz et al., 2012). Therefore, weakness in central coherence entails difficulties in contextual comprehension of social situations and in adaption to these situations (López and Leekam, 2007). Difficulties in central coherence are also described by (Lamb and Robertson, 1989) as local bias and by (Vermeulen, 2015), in his studies on people with autism spectrum disorders (ASD), as context blindness. To date, dysfunction in central coherence has principally been studied in ASD (López and Leekam, 2007), although it has also been described in other disorders, such as, Down's syndrome, Williams syndrome (D'Souza et al., 2016) non-verbal learning disorder (NVLD) (Gillberg, 2003, 2009; Crespo-Eguílaz and Narbona, 2011); and thus weak central coherence is apparently not specific to ASD.

1.1 Central coherence in autism spectrum disorder

It has been demonstrated that people with ASD tend to process information in a manner that is more focussed on details than on the overall meaning, that is, they have weak central coherence (Frith, 1992; Nydén et al., 2011). Theory concerning central coherence is supported by studies that confirm that, in tests of local preference, people with ASD perform significantly better than people of normotypical development (Shah and Frith, 1983; Jolliffe and Baron-Cohen, 1997; Plaisted et al., 1998; O'Riordan et al., 2001; Pellicano et al., 2006), which demonstrates that people with ASD tend to focus their attention more on parts of objects than on the objects themselves (Ornitz et al., 1977). Other research has shown that subjects with ASD perform poorly in tests of global preference (Mottron and Belleville, 1993; Rinehart et al., 2000), that is, in tests that include tasks that require detection of relatively small visual elements embedded in large fields (Caron et al., 2006), of visual searching (O'Riordan et al., 2001), of pattern discrimination (Plaisted et al., 1998; Bertone et al., 2005) or that involve design of blocks, impossible figures or embedded figures (Happé and Frith, 2006). Some studies find that in ASD subjects there is hyperfunction of brain areas generally involved in primary perception in contrast to perceptual integration, and researchers have proposed that this hyperfunction might be the explanation for the perceptual endophenotype in autism. The special abilities of the so-called autistic savant and the variability across the spectrum of ASD are possible manifestations of a tendency to use primary perceptual functions (Mottron et al., 2006). Various studies of people with "high-functioning autism" have led to conclusions compatible with this hypothesis (Bertone et al., 2003; Wang et al., 2007; Koldewyn et al., 2013; Syriopoulou Delli et al., 2016): participants with autism perform well at tasks involving perception of faces in stationary images (photos of faces) (Lahaie et al., 2006), at tasks requiring perception of movement (Bertone et al., 2004) and at the Wechsler scale Cubes test (Caron et al., 2006).

A modified version of the central coherence theory has been proposed in which it is hypothesised that in individuals with ASD the bias towards local processing can be overcome when doing tasks with explicit demands for global processing (Happé and Frith, 2006). According to this model, people with ASD do not necessarily have difficulties in perceiving the global form of things but rather have an over-specialized perceptual system that, depending on the requirements of a task, can interfere with higher-level cognition (Mottron and Burack, 2001; Caron et al., 2006).

Finally, despite otherwise contradictory findings (Happé, 1999), it has been established that the local-precedence style of information processing is not universally present across the whole autism spectrum (Happé, 1999).

1.2 Central coherence in non-verbal learning disorder

Children with non-verbal learning disorder (NVLD)—also denominated Deficits in Attention, Motor control and Perception (DAMP) and Procedural Learning Disorder (PLD) (Crespo-Eguílaz and Narbona, 2009)—show signs of weak central coherence. Such children can get lost in details rather than process information in an integrated and correct way (Doty, 2019). They find it difficult or are slow to arrive at a coherent comprehension of complex images or scenarios (Crespo-Eguílaz and Narbona, 2009; Magallón, 2011; Crespo-Eguílaz et al., 2012). They tend not to understand globally but rather in parts, which makes it difficult for them to carry out integration of concepts and abstraction and, therefore, to make correct adaptation of understanding to context (Díaz Lucero et al., 2011); they have difficulty perceiving globally, analyzing, organizing and summarizing information (Chow and Skuy, 1999; Molenaar-Klumper, 2002; Mammarella and Pazzaglia, 2010). In addition, it has been found that these children can perform poorly in certain visual perception tasks (for example, in perceiving the spatial location of objects), and have difficulties in recognizing what is detail, simultaneous processing, combining parts into a whole, and visual-spatial organization (Schoemaker et al., 2001; Drumond et al., 2005). Sometimes, children with NVLD make errors in spatial perception, for example, mistaking places in their surroundings or in the position of a person relative to themselves (Viñuela, 2007). Consequently, they find it difficult to cope with novel environments and to solve problems that have a visual-spatial component (González, 2017). They tend to get lost in open, unstructured situations in which conversations often overlap, there is more use of colloquial language, many gestures are used and body distances need to be managed (Foss, 2001). They evidence both deficits in the comprehension of extra-verbal information (facial expressions, gestures, mimicry, body postures, prosodic inflections and other visual aspects of their circumstances) and also difficulties in integrating and understanding such information (Humphries et al., 2004; Mammarella et al., 2009). They find it challenging to adapt to novel situations and tend to make generalisations based on specific verbal utterances, without taking into account the context in which a conversation is taking place (Worling et al., 1999).

The above-mentioned impediments have a major impact on people with NVLD when it comes to giving meaning to different contexts and interpreting discourse and affects the contextualisation of language. They find it difficult to understand figurative language, irony and jokes; they may interpret language literally and have problems adapting to novel situations of social interaction (Semrud-Clikeman and Hynd, 1990; Colomé Roura et al., 2009; Narbona et al., 2011). Consequently, people with NVLD are unable to communicate effectively in everyday situations (Colomé Roura et al., 2009) and experience difficulties in social relations.

1.3 Central coherence in attention deficit and hyperactivity disorder

Booth and Happé (Booth and Happé, 2010) compared how well participants with ADHD, with ASD and with normotypical development performed at completing sentences or phrases (for example, "Hunting with a knife and . . . fork"); they found that the participants with ADHD correctly performed at this task, while the ASD group performed significantly worse than the control group.

In a study by Crespo-Eguílaz et al. (2012), 200 school children—20 with NVLD, 60 with ADHD, 60 with both NVLD and ADHD, and 60 controls - were given a test involving a chimerical image and an incoherent visual scene. Of the children with ADHD, only 8% failed in rapid interpretation of the chimerical image, and only 7% performed poorly in comprehension of the visual scene. In a similar research (Magallón, 2011), found that only 13% of schoolchildren with ADHD performed badly in the chimerical image task. Similar findings are founded by different authors (Booth et al., 2003; Zhang and Adipat, 2005; Pina et al., 2013).

1.4 Central coherence in social communication disorder

Children with SCD tend to interpret language literally and not to detect irony, inference and/or metaphors (Bishop and Rosenbloom, 1987; Bishop and Adams, 1992; Leinonen and Letts, 1997; Bishop, 2000; Mulas Delgado et al., 2006; Velarde et al., 2017). They present problems in adapting their language to the needs of the listener or the situation, and they also lack flexibility when topics change during a dialogue (Rapin and Allen, 1983; McTear, 1985; Conti-Ramsden and Gunn, 1986; Bishop, 2000; Mendoza Lara and Muñoz López, 2005; González et al., 2015).

In the DSM-5 criteria for SCD (American Psychiatric Association [APA], 2013), defining characteristics of the disorder are deficiency in the use of communication for social purposes (such as, greeting people and sharing information in a manner appropriate to the social context or to the needs of the person listening) and difficulty understanding what is not said in an explicit manner and what is non-literal or ambiguous (for example, idioms and humor) (Monfort, 2001; Baixauli-Fortea et al., 2004; Perkins, 2010; Monfort Juárez Centro Entender Hablar I, 2013; Martínez Alonso et al., 2021).

Given this review of the literature, our objective is to confirm that schoolchildren with ASD and NVLD experience difficulties with central coherence. Additionally, we aim to test whether children with SCD exhibit weak central coherence, a hypothesis suggested by our clinical observations but not yet empirically verified. Furthermore, we will examine the performance of children with ADHD concerning this construct and determine whether the central coherence deficit observed in children with NVLD can be explained by their attention difficulties and/or hyperactivity.

2 Materials and methods

2.1 Participants

The study sample was of 252 participants comprising a control group of 194 normotypical schoolchildren and 58 children with clinical disorders recruited at the neuropediatrics unit of the Clínica Universidad de Navarra hospital. The disorders were non-verbal learning disorder in conjunction with attention deficit and hyperactivity disorder (NVLD + ADHD; n = 24); ADHD (n = 16); level 1 autism spectrum disorder (ASD1 = 10) and social communication disorder (SCD; n = 8). All children were at primary school, between 6 and 11 years old (**Table 1**), and had typical intelligence as evaluated by Raven's Progressive Matrices Test (2001) (**Table 1**). The children were from families of middle to middle-high socio-economic and cultural level on the Hollingshead scale (Hollingshead, 1957). All the children were Caucasian in race.

The proportions of boys and girls and IQ-related statistics are given in **Table 1**. IQ data was not available for the control group; exclusion criteria for the control group included low academic performance, learning difficulty or behavioural disorders as determined, at the time of the study, by teachers and other specialist education professionals. 45.8% of the participants with NVLD+ADHD and 62.5% of those with ADHD were receiving pharmacological treatment with methylphenidate to improve attention (**Table 4**). The questionnaire and methodology for this study was approved by the Human Research Ethics committee of the University of Navarra (Ethics approval number: 2017.004mod1).

Written informed consent was obtained from the parents (**Tables 1**, **2**).

2.2 Tools and procedures

All children took the Central Coherence Test (CCT) (Gambra, 2020), which is an in-house development comprised of 36 items grouped in four dimensions, each of which has, in turn, different visual and verbal sub-tests or tasks (**Table 3**).

Previous research has demonstrated the validity and reliability of the CCT (Gambra, 2020) by means of reliability analysis (Cronbach's alpha and Spearman's rank correlation coefficient) and establishing construct validity (exploratory and confirmatory factor analysis and other studies), and establishing convergent and discriminant validity.

The dimensions that make up the test, as well as the subtests of each dimension, are as follows:

TABLE 1 Sex distribution and IQ-related statistics for the study sample.

	Control group	Clinical sample			
		NVLD+ADHD	ADHD	SCD	ASD1
n	194	24	16	8	10
Male	76	22	14	4	8
Female	118	2	2	4	2
Ratio male: female	0.6:1	11:1	7:1	1:1	4:1
IQ: mean (SD)	–	100.4 (9.9)	104.8 (8.7)	101.7 (11.2)	100 (12.4)
IQ: min.-max.	–	82–129	83–114	83–119	87–119

The above data cannot be used to infer prevalence because the subjects in this study volunteered to participate: not all patients with these pathologies who attended our neuropediatrics unit chose to participate.

TABLE 2 Age distribution of the sample studied.

Age	Control group	Clinical sample	Clinical sample by pathology			
			NVLD+ADHD	ADHD	SCD	ASD1
6	36	5	2	0	1	2
7	39	9	2	4	1	2
8	40	8	4	2	1	1
9	16	11	4	6	0	1
10	38	7	3	2	1	1
11	25	18	9	2	4	3
Total	194	58	24	16	8	10

(a) Inference: The tasks that make up this dimension assess the ability to give meaning to different sentences and texts according to the contextual situation in which they exist, as well as the ability to understand figurative language, irony and jokes.

  a. Irony: assesses the understanding of various ironic comments made in three different everyday situations.
  b. Literality: assesses the ability to understand the non-literal meaning of a series of sentences on the basis of the context in which each sentence occurs.
  c. Verbal story-telling: assesses ability to provide a coherent ending to incomplete stories.

(b) Verbal detail: this dimension evaluates ability to detect incoherent features in different situations presented verbally.

  a. Nonsense questions: assesses ability to detect coherence or incoherence in a series of nonsense questions.
  b. Nonsense sentences: assesses ability to detect inconsistency in sentences that are grammatically correct but that are inconsistent in terms of content. The nonsense sentences are mixed in with a series of sensible sentences, which serve as distractors.

(c) Simultaneity: this dimension assesses the ability and speed of the schoolchild in making sense of an inconsistent situation: a series of non-coherent images and texts.

  a. Chimerical images: these evaluate simultanagnosia. After seeing each chimeric image for two seconds, the child is asked whether he/she has identified both, one or none of the animals or objects that make up the chimera. The child is also asked to describe the inconsistency between the two components.
  b. Inconsistent pictures: these are used to evaluate ability to orally describe illustrated actions that are inherently inconsistent.
  c. Hidden numbers and colours: this test evaluates ability to perceive details (numbers and colours) within a whole (a story). The test is done both visually (the child sees and reads the story) and aurally (the child hears the story).

(d) Context: the subtests in this dimension evaluate ability to understand and freely describe, orally, a series of images and audio recordings in which various actions are represented.

  a. Inconsistent visual scenes: these evaluate ability to detect incongruities in various images.
  b. The phone call: this is a series of role-play situations to assess ability to adapt to different contexts (Table 3).

We used Kolmogorov-Smirnov and Shapiro-Wilk tests to determine whether variables had a normal distribution. In addition Levene's test was used to assess whether variances for the different groups were equal. For each clinical group, values for each test variable were converted to typified scores (for subtests and dimensions of the CCT, and for the CCT as a whole). For each group the performance profile was prepared. The non-parametric Kluskall-Wallis H and Mann-Whitney U were calculated to make comparisons within groups and between pairs of groups.

TABLE 3 Dimensions and subtests of the Central Coherence Test.

Dimensions	Subtests	No of items	Items
Inference		11	1–11
	Irony	3	1–3
	Literality	7	4–10
	Verbal story-telling	1	11
Verbal detail		6	12–17
	Nonsense questions	2	12–13
	Nonsense sentences	4	14–17
Simultaneity		9	18–26
	Chimerical images	4	18–21
	Inconsistent pictures	1	22
	Hidden numbers and colours	4	23–26
Context		10	27–36
	Inconsistent visual scenes	4	27–30
	The phone call	6	31–36
Total		36	

Also, in order to determine whether medication influences test performance, we used Student's t test to compare the mean performance of participants under medication with the mean performance of those who were not.

TABLE 4 Influence of medication on performance in dimensions and tasks of the Central Coherence Test.

	NVLD + ADHD group	ADHD group
Not under medication (*n*)	13	6
Under medication (*n*)	11	10
Dimension/task of the Central Coherence Test:	Student's *t* comparing mean performance of medicated and non-medicated children:	
Inference	0.352	0.310
Irony	0.247	0.480
Literality	0.811	0.628
Verbal story-telling	0.397	0.191
Verbal detail	0.756	0.051
Nonsense questions	0.894	0.610
Nonsense sentences	0.283	0.272
Simultaneity	0.519	0.595
Chimerical images	0.416	0.063
Inconsistent pictures	0.827	0.319
Hidden numbers and colours	0.436	0.484
Context	0.950	0.911
Inconsistent visual scenes	0.550	0.665
The phone call	0.168	0.262

3 Results

There were no significant differences in performance between medicated and non-medicated participants. This was the case for all dimensions of the CCT and for all tasks. The finding applies to the two clinical groups in which there were patients under pharmacological treatment: the NVLD+ADHD group and the ADHD group (Table 4).

There were no significant differences between clinical groups or the control group in performance in the Verbal detail dimension.

With regard to the Simultaneity dimension, the performance levels of the NVLD+ADHD group and the SCD group were significantly poorer than that of the control group (Figure 1); the effect size is high (Table 5). Performance levels of the ADHD and ASD1 groups were typical (Figure 1). There were significant differences in this respect between the NVLD+ADHD group and the ADHD and ASD1 groups (Figure 1; Table 5).

The performance of NVLD+ADHD and ASD1 groups in the Inference dimension was significantly lower than that of the control group, while the performance of the ADHD group was typical. The effect size of the differences was high (Table 5). The performance of the SCD group was also apparently lower, but this difference was not statistically significant (Figure 1).

In the Context dimension, the mean performance levels of the NVLD+ADHD group and the ASD1 group were significantly lower than that of the control group (Figure 1 and Table 5). Mean performance of subjects in ADHD and SCD groups was adequate. The effect size was high for all comparisons.

Finally, for the CCT as a whole, mean performance levels of the NVLD+ADHD group and of the SCD group were significantly lower (with large effect size) than that of the control group, while mean performance levels of the ADHD and ASD1 groups were average (Figure 1 and Table 5). Overall mean performance of the NVLD+ADHD group was significantly lower than that of the ADHD group (Table 5).

In total CCT score and in all dimensions except Verbal detail, the NVLD+ADHD group performed significantly worse than the control group. The performance of the ADHD group, however, was similar to that of the control group in all dimensions. Furthermore, the NVLD+ADHD and ADHD groups differed in overall CCT (Table 6). As the participants in both groups have attention deficit, we infer that the clear difficulties in central coherence in subjects with NVLD+ADHD is not to be explained by attention deficit but rather as being characteristic of NVLD.

NVLD+ADHD and ASD1 groups had significantly lower mean performance than the control group in the Inference and Context dimensions (Table 6). The two groups differ in mean performance in Simultaneity: children with NVLD+ADHD performing worse than those with ASD1.

CCT performance profiles for NVLD+ADHD and SCD groups were similar. In both groups, mean performance in Simultaneity was significantly lower than that of the control group. However, the performance deficit of the NVLD+ADHD group was more serious than that of the SCD group because the symptomatology was more

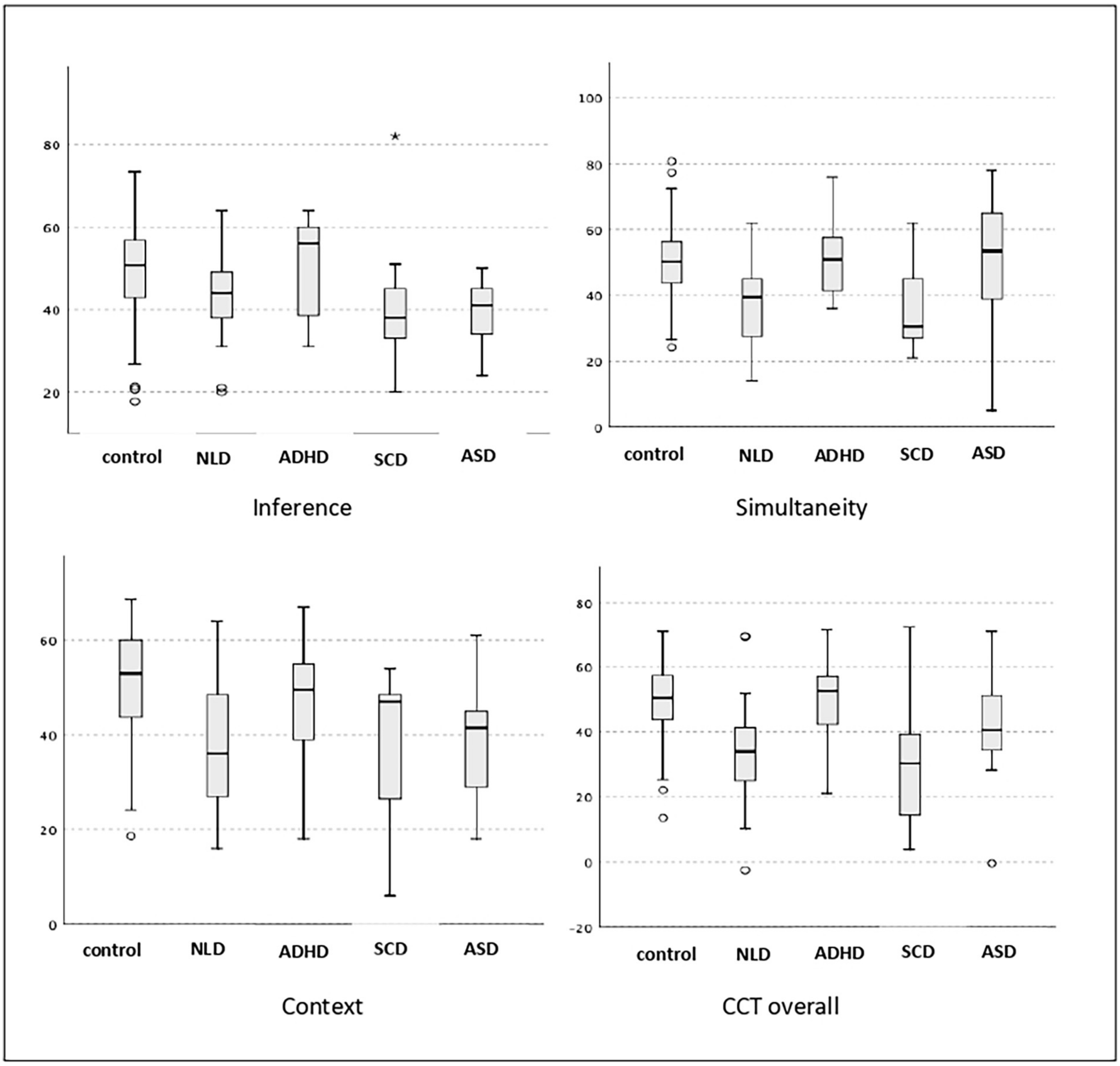


FIGURE 1
Comparison of performance levels of the clinical groups in the dimensions of the CCT and in the CCT as a whole.

pronounced: the mean performance of the NVLD+ADHD group was significantly lower than that of the control group in most dimensions of the CCT (**Table 6**).

4 Discussion

The results of our study indicate that deficit in central coherence is not an effect exclusive to ASD. While various studies have found a reduction in global processing in ASD (Rinehart et al., 2000; Pellicano et al., 2005; Seernani et al., 2020) or that people with ASD perform significantly better in tests of local preference than people with typical development (Shah and Frith, 1983; Plaisted et al., 1998; O'Riordan et al., 2001; Pellicano et al., 2006; Gambra, 2020), there are also numerous studies that contradict such findings (Brian and Bryson, 1996) and that demonstrate that ability for global processing in people with ASD is intact (Bertone et al., 2003; Wang et al., 2004; Johnson et al., 2010).

Mottron and Belleville (Mottron and Belleville, 1993) demonstrate that people with autism process information, at both the local and the global level, as well as control subjects, but that in ASD the local interferes with the global when stimuli are incongruous. Subsequent studies confirm these findings (Jolliffe and Baron-Cohen, 1997; Rinehart et al., 2001). These somewhat contradictory findings have also been obtained in studies using more than one evaluation tool (Edgin and Pennington, 2005). Syriopoulou Delli et al. (2016) suggest that disparity in results might be best explained by considering that the style of information processing in autism is personal rather than a distinct characteristic of ASD.

In our research, the general performance of the ASD1 group was lower than that of the control group, but this difference was not always significant (depending on the dimension of the

TABLE 5 Comparisons of pairs of groups in performance in the dimensions of Inference, Simultaneity and Context and in overall CCT performance.

	χ	Sig.	Cohen's D
Inference			
Control *vs.* NVLD+ADHD	46.52	**0.032**	6.79
Control *vs.* ADHD	−7.48	1	−0.56
Control *vs.* SCD	64.14	0.147	8.25
Control *vs.* ASD1	73.09	**0.020**	10.10
NVLD+ADHD *vs.* SCD	17.64	1	-
NVLD+ADHD *vs.* ASD1	26.58	1	-
NVLD+ADHD *vs.* ADHD	−54.00	0.217	-
SCD *vs.* ADHD	71.63	0.232	-
SCD *vs.* ASD1	8.95	1	-
ADHD *vs.* ASD1	80.58	0.061	-
Simultaneity			
Control *vs.* NVLD+ADHD	62.64	**0.001**	12
Control *vs.* ADHD	−5.32	1	−1.19
Control *vs.* SCD	74.8	**0.044**	14
Control *vs.* ASD1	−14.47	1	−0.30
NVLD+ADHD *vs.*SCD	12.17	1	-
NVLD+ADHD *vs.*ASD1	−77.11	**0.049**	-
NVLD+ADHD *vs.* ADHD	−67.96	0.**039**	-
SCD *vs.* ADHD	80.13	0.111	-
SCD *vs.* ASD1	−89.28	0.098	-
ADHD *vs.* ASD1	−9.15	1	-
Context			
Control *vs.* NVLD+ADHD	69.81	< **0.001**	13.08
Control *vs.* ADHD	22.16	1	3.93
Control *vs.* SCD	59.13	0.245	12.75
Control *vs.* ASD1	66.31	**0.049**	11.34
NVLD+ADHD *vs.* SCD	−10.69	1	-
NVLD+ADHD *vs.* ASD1	−3.50	1	-
NVLD+ADHD *vs.* ADHD	−47.66	0.391	-
SCD *vs.* ADHD	36.97	1	-
SCD *vs.* ASD1	7.19	1	-
ADHD *vs.* ASD1	44.16	1	-
Total CCT			
Control *vs.* NVLD+ADHD	79.14	< **0.001**	15.89
Control *vs.* ADHD	−3.10	1	−0.21
Control *vs.* SCD	84.80	**0.013**	19.36
Control *vs.* ASD1	44.24	0.622	9.39
NVLD+ADHD *vs.* SCD	5.67	1	-
NVLD+ADHD *vs.* ASD1	−34.90	1	-
NVLD+ADHD *vs.* ADHD	−82.24	**0.005**	-
SCD *vs.* ADHD	87.90	0.55	-
SCD *vs.* ASD1	−40.56	1	-
ADHD *vs.* ASD1	47.34	1	-

Significance values less than or equal to 0.05 are considered to be statistically significant. The Bonferroni correction for multiple-comparison testing has been applied to significance values. Effect sizes are interpreted as follows: $d < 0.20$: small; $0.20 < d < 0.80$: average; $d > 0.80$: large. Bold values indicate statistically significant.

TABLE 6 CCT performance profiles based on differences with respect to the control group.

	Relative to the control group		
	NVLD+ADHD	ASD1	SCD
Inference	average-low	average-low	average-low
Verbal detail	average	average-low	average-low
Simultaneity*	average-low	average	average-low
Context	average-low	average-low	average-low
CCT Total	low	average-low	average

A gray background indicates differences for which $p < 0.001$. *In Simultaneity there was a statistically significant difference ($p = 0.049$) between NVLD+ADHD and ASD1 groups.

CCT); this is in agreement with the results of Syriopoulou Delli et al. (2016) who found that in a global preference task the scores of children with typical development were apparently higher than the scores of children with ASD, but the difference was not statistically significant. With regard to central coherence abilities, schoolchildren with ASD1 had average performance in simultaneity tasks (which were principally visual tasks) and average-low performance in tasks involving inference, verbal detail and understanding in context. Relative to the control group, differences in performance at tasks involving inference and context were significant. (The same differences were found in the NVLD+ADHD group, but children with ASD1 performed better than children with NVLD+ADHD in simultaneity tasks.)

In contrast to our findings, Loth (Loth, 2003) concluded that only about 35% of children with autism show weak central coherence in different tasks, while 48% have mixed styles of processing, with good performance in conceptual tasks. In our study, however, 90% of the schoolchildren in the ASD1 group had difficulty with one or more abilities related to central coherence. The difference in results can be explained in various ways. First, ASD is by nature heterogeneous and symptoms vary according to the level of ASD (Santangelo and Folstein, 1999; Tager-Flusberg and Joseph, 2003); in our study, all participants with ASD were level 1. Second, within ASD there can be subgroups defined by performance in terms of central coherence (Tager-Flusberg and Joseph, 2003). Third, as already discussed, the different tasks used by different studies make different demands on subjects and therefore give different measures of performance. Finally, Mottron et al. (Mottron et al., 2006) suggest that autism is often associated with improved perceptual processing, but this is not evident in all children with ASD.

Our study found that children with NVLD had difficulties with central coherence, confirming that this difficulty is a characteristic in the cognitive profile of NVLD. This finding is in line with different studies (Crespo-Eguílaz and Narbona, 2009; Magallón, 2011). Other authors (Chow and Skuy, 1999; Schoemaker et al., 2001; Molenaar-Klumper, 2002; Viñuela, 2007; Mammarella and Pazzaglia, 2010; Mammarella et al., 2019) do not refer explicitly to the construct of central coherence but nonetheless affirm that children with NVLD have difficulty with global perception and with analysing, organizing and synthesizing information.

General performance of the NVLD+ADHD group in central coherence tasks was lower than that of the control group ($p < 0.001$). However, performance was average for abilities related to verbal details. That is, for children with NVLD+ADHD, performance in tasks that involve verbal aspects of central coherence seems to stand out against the background of their overall performance. This concurs with the more general observation that the verbal abilities of children with NVLD+ADHD are better than their manipulative abilities. In our study, performance was average-low for abilities to understand in a simultaneous manner, to make inferences and to understand in a context (all of which were statistically significant differences relative to the control group). These results are coherent with those of Crespo-Eguílaz et al. (2012), who reported that rapid interpretation of a chimerical image - to make sense of which it was necessary to perceive and integrate both parts of the chimera - posed difficulty for 85% of schoolchildren with NVLD+ADHD but for only 5% of controls. Magallón (2011) found that children with NVLD+ADHD had difficulty in noticing incongruities in chimerical images (62.1% carried out this task badly in comparison to 8% of controls) and in visual scenes (about 60% performed poorly at this task). These results are along the same lines (Díaz Lucero et al., 2011), who evaluated the neuropsychological profile of a group of 22 children with DAMP and concluded that they did not understand globally but rather by parts, which made it difficult for them to integrate concepts, to carry out abstraction and, therefore, to adapt correctly to the context. Another research group, (Drumond et al., 2005), also found that children with NVLD had difficulty in tasks that involved construction of a whole from parts. Semrud-Clikeman et al. (2010) reported findings similar to ours for a group of children with NVLD (with difficulties in central coherence) relative to other groups of children with level-1 ASD and ADHD (without difficulties in central coherence).

In our study, the children with NVLD also had ADHD. However, as in the study of Makris et al. (2021), the difficulties in central coherence that we have discussed were not observed with a separate ADHD group. Therefore, these difficulties cannot be explained by attention deficit and appear to be characteristic of NVLD. In contrast to these results, (Cardillo, 2018) found that the central coherence profile of a group of children with ADHD was heterogeneous: the children had difficulty with visuoconstruction abilities when they had to battle with global configurations but performed visual-perception tasks correctly. In this study, the SCD group was characterized by average-low performance for all central coherence abilities studied. Thus, as a group, schoolchildren with a deficit in social communication skills have below-average performance, which indicates that their neuropsychological profile is also characterized by weak central coherence.

The study's limitations include that all children in the NVLD sample have ADHD. It would have been preferable to identify another pure NVLD group, although this is challenging due to the frequent presence of comorbid disorders. Additionally, the sample size of the ASD group could be larger to draw definitive conclusions. Environmental or socio-economic factors not accounted for in the study could influence central coherence abilities. These factors might limit the generalizability of findings to different populations.

5 Conclusion

The current study, through the profiles of central coherence for the clinical groups described, is consistent with a deficit in central coherence is not exclusive to autism spectrum disorders. This study evidence that that children with other neurodevelopmental and learning disorders, such as, non-verbal learning disorder and social communication disorder, experience difficulties related with this cognitive function. In addition, it was found that schoolchildren with ADHD did not have difficulty with central coherence. Finally, we establish that the Central Coherence Test provides complementary information that is useful for differential diagnosis between neurodevelopmental disorders involving weak central coherence.

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 the Ethics committee of the University of Navarra (Ethics approval number: 2017.004mod1). 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. Written informed consent was obtained from the minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

Author contributions

LG: Writing–original draft, Conceptualization, Methodology, Writing–review and editing. SM: Formal analysis, Writing–original draft, Writing–review and editing. NC-E: Formal analysis, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of the article. This study was funded by SM Grant PID2020-119328GA-I00 funded by MCIN/AEI/ 10.13039/501100011033.

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

American Psychiatric Association [APA] (2013). *Diagnostic and statistical manual of mental disorders DSM-5*. Washington, DC: American Psychiatric Association [APA]. doi: 10.1176/appi.books.9780890425596

Attwood, T. (2007). *Guía del síndrome de Asperger*. Barcelona: Paidòis.

Baixauli-Fortea, I., Roselló, B., and Miranda-Casas (2004). Evaluación de las dificultades pragmáticas. Estudio de casos. *Rev. Neurol.* 38, 69–79. doi: 10.33588/rn.38S1.2004063

Bertone, A., Mottron, L., and Faubert, J. (2004). Autism and schizophrenia: Similar perceptual consequence, different neurobiological etiology? *Behav. Brain Sci.* 27, 592–593. doi: 10.1017/S0140525X04260137

Bertone, A., Mottron, L., Jelenic, P., and Faubert, J. (2003). Motion perception in autism: A "complex" Issue. *J. Cogn. Neurosci.* 15, 218–225. doi: 10.1162/089892903321208150

Bertone, A., Mottron, L., Jelenic, P., and Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain* 128, 2430–2441. doi: 10.1093/brain/awh561

Bishop, D. V. M. (2000). How does the brain learn language? Insights from the study of children with and without language impairment. *Dev. Med. Child Neurol.* 42, 133–142. doi: 10.1017/S0012162200000244

Bishop, D. V. M., and Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential meaning. *J. Speech Hear. Res.* 35, 119–129. doi: 10.1044/jshr.3501.119

Bishop, D. V., and Rosenbloom, L. (1987). Classification of childhood language disorders. *Lang. Dev. Disord.* 22, 61–81.

Booth, R., and Happé, F. (2010). "Hunting with a knife and fork": Examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic task. *J. Exp. Child Psychol.* 107, 377–393. doi: 10.1016/j.jecp.2010.06.003

Booth, R., Charlton, R., Hughes, C., and Happé, F. (2003). Disentangling weak coherence and executive dysfunction: Planning drawing in autism and attention–deficit/hyperactivity disorder. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 387–392. doi: 10.1098/rstb.2002.1204

Brian, J. A., and Bryson, S. E. (1996). Disembedding performance and recognition memory in autism/PDD. *J. Child Psychol. Psychiatry* 37, 865–872. doi: 10.1111/j.1469-7610.1996.tb01482.x

Cardillo, R. (2018). *Local-global visuospatial processing in autism spectrum disorders and nonverbal learning disabilities: A cross-task and cross-disorder comparison*. Padua: University of Padova.

Caron, M. J., Mottron, L., Berthiaume, C., and Dawson, M. (2006). Cognitive mechanisms, specificity and neural underpinnings of visuospatial peaks in autism. *Brain* 129, 1789–1802. doi: 10.1093/brain/awl072

Chow, D., and Skuy, M. (1999). Simultaneous and successive cognitive processing in children with nonverbal learning disabilities. *Sch. Psychol. Int.* 20, 219–231. doi: 10.1177/0143034399202005

Colomé Roura, R., Sans Fitó, A., López Sala, A., and Boix Lluch, C. (2009). Trastorno de aprendizaje no verbal: Características cognitivo-conductuales y aspectos neuropsicológicos. *Rev. Neurol.* 48:77. doi: 10.33588/rn.48S02.2009009

Conti-Ramsden, G., and Gunn, M. (1986). The development of conversational disability: A case study. *Br. J. Disord. Commun.* 21, 339–351. doi: 10.3109/13682828609019846

Crespo-Eguílaz, N., and Narbona, J. (2009). Trastorno de aprendizaje procedimental: Características neuropsicológicas. *Rev. Neurol.* 49:409. doi: 10.33588/rn.4908.2009079

Crespo-Eguílaz, N., and Narbona, J. (2011). Dificultades en la percatación ràipida de incongruencias en el trastorno de aprendizaje procedimental: Posible disfunción de la coherencia central. *Rev. Neurol.* 52(Suppl. 1), 39–41. doi: 10.33588/rn.52S01.2010807

Crespo-Eguílaz, N., Narbona, J., and Magallón, S. (2012). Disfunción de la coherencia central en niños con trastorno de aprendizaje procedimental. *Rev. Neurol.* 55:513. doi: 10.33588/rn.5509.2012334

D'Souza, D., Booth, R., Connolly, M., Happé, F., and Karmiloff-Smith, A. (2016). Rethinking the concepts of 'local or global processors': Evidence from Williams syndrome, down syndrome, and autism spectrum disorders. *Dev. Sci.* 19, 452–468. doi: 10.1111/desc.12312

Díaz Lucero, A. H., Melano, C. A., and Etchepareborda Simonini, M. C. (2011). Síndrome de déficit de atención, del control motor y de la percepción (DAMP): Perfil neuropsicológico. *Rev. Neurol.* 52:S071. doi: 10.33588/rn.52S01.2010797

Doty, N. (2019). "The Massachusetts general hospital guide to learning disabilities," in *Current clinical psychiatry*, eds H. K. Wilson and E. B. Braaten (Cham: Humana Press), doi: 10.1007/978-3-319-98643-2

Drumond, C., Ahmad, S., and Rourke, B. (2005). Rules for the classification of younger children with nonverbal learning disabilities and basic phonological processing disabilities. *Arch. Clin. Neuropsychol.* 20, 171–182. doi: 10.1016/j.acn.2004.05.001

Edgin, J. O., and Pennington, B. F. (2005). Spatial cognition in autism spectrum disorders: Superior, impaired, or just intact? *J. Autism Dev. Disord.* 35, 729–745. doi: 10.1007/s10803-005-0020-y

Förster, J., and Dannenberg, L. (2010). GLOMO sys: A systems account of global versus local processing. *Psychol. Inq.* 21, 175–197. doi: 10.1080/1047840X.2010.487849

Foss, J. M. (2001). "Nonverbal learning disability: How to recognize it and minimize its effects," in *Proceedings of the ERIC clearinghouse on disabilities and gifted education, council for exceptional children*, (Arlington, VA).

Frith, U. (1992). *Autismo: Hacia una explicación del enigma*. madrid: Alianza Editorial.

Gambra, L. (2020). *Central coherence: Design and validation of an assessment and study test in procedural learning disorder.* Pamplona: University of Navarra.

Gillberg, C. (2003). Deficits in attention, motor control, and perception: A brief review. *Arch. Dis. Child.* 88, 904–910. doi: 10.1136/adc.88.10.904

Gillberg, C. (2009). "Developmental and neuropsychiatric disorders of childhood," in *Diseases of the nervous system in childhood*, 3rd Edn, eds J. Aicardi, M. Bax, and C. Gillberg (London: Mac Keith Press), 889–901.

González, D. (2017). Un trastorno desconocido: Trastorno de aprendizaje no verbal. *Padres Maestros* 369, 17–19. doi: 10.14422/pym.i369.y2017.002

González, M., Rivas, R. M., and López, S. (2015). Caracterización y delimitación del trastorno de la comunicación social (pragmático). *Rev. Estud. Invest. Psicol. Educ.* 24, 005–8. doi: 10.17979/reipe.2015.0.09.132

Happé, F. (1999). Autism: Cognitive deficit or cognitive style? *Trends Cogn. Sci.* 3, 216–222. doi: 10.1016/S1364-6613(99)01318-2

Happé, F., and Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25. doi: 10.1007/s10803-005-0039-0

Hollingshead, A. B. (1957). *Two factors index of social position*. New Haven, CT: Yale University.

Humphries, T., Cardy, J. O., Worling, D. E., and Peets, K. (2004). Narrative comprehension and retelling abilities of children with nonverbal learning disabilities. *Brain Cogn.* 56, 77–88. doi: 10.1016/j.bandc.2004.06.001

Johnson, S. A., Blaha, L. M., Houpt, J. W., and Townsend, J. T. (2010). Systems factorial technology provides new insights on global–local information processing in autism spectrum disorders. *J. Math. Psychol.* 54, 53–72. doi: 10.1016/j.jmp.2009.06.006

Jolliffe, T., and Baron-Cohen, S. (1997). Are people with autism and asperger syndrome faster than normal on the embedded figures test? *J. Child Psychol. Psychiatry* 38, 527–534. doi: 10.1111/j.1469-7610.1997.tb01539.x

Koldewyn, K., Jiang, Y. V., Weigelt, S., and Kanwisher, N. (2013). Global/local processing in autism: Not a disability, but a disinclination. *J. Autism Dev. Disord.* 43, 2329–2340. doi: 10.1007/s10803-013-1777-z

Lahaie, A., Mottron, L., Arguin, M., Berthiaume, C., Jemel, B., and Saumier, D. (2006). Face perception in high-functioning autistic adults: Evidence for superior processing of face parts, not for a configural face-processing deficit. *Neuropsychology* 20, 30–41. doi: 10.1037/0894-4105.20.1.30

Lamb, M. R., and Robertson, L. C. (1989). Do response time advantage and interference reflect the order of processing of global- and local-level information? *Percept. Psychophys.* 46, 254–258. doi: 10.3758/BF03208087

Leinonen, E., and Letts, C. (1997). Referential communication tasks: Performance by normal and pragmatically impaired children. *Int. J. Lang. Commun. Disord.* 32, 53–65. doi: 10.1111/j.1460-6984.1997.tb01624.x

López, B., and Leekam, S. R. (2007). Teoría de la coherencia central: Una revisión de los supuestos teóricos. *Inf. Aprend.* 30, 439–457. doi: 10.1174/021037007781787462

Loth, E. (2003). *On social, cultural and cognitive aspects of theory of mind in practice (ProQuest No. 10166930)*. St Andrews: University of St Andrews.

Magallón, S. (2011). *Memoria procedimental en escolares típicos y en escolares con trastornos de aprendizaje*. Pamplona: Universidad de Navarra.

Makris, G., Pervanidou, P., Chouliaras, G., Stachtea, X., Valavani, E., Bastaki, D., et al. (2021). Diverse patterns of vulnerability to visual illusions in children with neurodevelopmental disorders. *Cogn. Process.* 22, 659–673. doi: 10.1007/s10339-021-01041-6

Mammarella, I. C., and Pazzaglia, F. (2010). Visual perception and memory impairments in children at risk of nonverbal learning disabilities. *Child Neuropsychol.* 16, 564–576. doi: 10.1080/09297049.2010.485125

Mammarella, I. C., Cardillo, R., and Zoccante, L. (2019). Differences in visuospatial processing in individuals with nonverbal learning disability or autism spectrum disorder without intellectual disability. *Neuropsychology* 33, 123–134. doi: 10.1037/neu0000492

Mammarella, I. C., Meneghetti, C., Pazzaglia, F., Gitti, F., Gomez, C., and Cornoldi, C. (2009). Representation of survey and route spatial descriptions in children with nonverbal (visuospatial) learning disabilities. *Brain Cogn.* 71, 173–179. doi: 10.1016/j.bandc.2009.05.003

Martínez Alonso, M., Fernández Rodríguez, M., Pérez Moleiro, L., and Belén Martínez Alonso, M. (2021). Trastorno de la comunicación social (pragmático), nueva categoría diagnóstica DSM-5, consideraciones clínicas y diagnóstico diferencial a propósito de un caso. *Rev. Psiquiatría Infanto Juvenil* 32, 155–160.

McTear, M. F. (1985). Pragmatic disorders: A case study of conversational disability. *Int. J. Lang. Commun. Disord.* 20, 129–142. doi: 10.3109/13682828509012255

Mendoza Lara, E., and Muñoz López, J. (2005). Del trastorno específico del lenguaje al autismo. *Rev. Neurol.* 41:S091. doi: 10.33588/rn.41S01.2005316

Molenaar-Klumper, M. (2002). *Non-verbal learning disabilities*. London: Jessica Kingsley Publishers.

Monfort Juárez Centro Entender Hablar I (2013). TRASTORNOS DEL LENGUAJE. *Rev. Neurol.* 56, S141–S146. doi: 10.33588/rn.56S01.2012667

Monfort, M. (2001). Niños con un déficit semántico-pragmático1. *Rev. Logop. Foniatría Audiol.* 21, 188–194. doi: 10.1016/S0214-4603(01)76208-0

Mottron, L., and Belleville, S. A. (1993). Study of perceptual analysis in a high-level autistic subject with exceptional graphic abilities. *Brain Cogn.* 23, 279–309. doi: 10.1006/brcg.1993.1060

Mottron, L., and Burack, J. A. (2001). "Enhanced perceptual functioning in the development of autism," in *The development of autism: Perspectives from theory and research*, eds J. A. Burack, T. Charman, N. Yirmiya, and R. Philip (Mahwah, NJ: Lawrence Erlbaum Associates), 131–148.

Mottron, L., Dawson, M., Soulières, I., Hubert, B., and Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *J. Autism Dev. Disord.* 36, 27–43. doi: 10.1007/s10803-005-0040-7

Mulas Delgado, F., Etchepareborda Simonini, M. C., Díaz Lucero, A. H., and Ruiz Andrés, R. (2006). El lenguaje y los trastornos del neurodesarrollo. Revisión de las características clínicas. *Rev. Neurol.* 42:S103. doi: 10.33588/rn.42S02.2005828

Narbona, J., Crespo-Eguílaz, N., and Magallón, S. (2011). "Trastorno de aprendizaje procedimental," in *Trastornos del neurodesarrollo*, eds J. Artigas-Pallarés and J. Narbona (Barcelona: Viguera Editores), 428–449.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cogn. Psychol.* 9, 353–383. doi: 10.1016/0010-0285(77)90012-3

Nayar, K., Voyles, A. C., Kiorpes, L., and di Martino, A. (2017). Global and local visual processing in autism: An objective assessment approach. *Autism Res.* 10, 1392–1404. doi: 10.1002/aur.1782

Noens, I. L. J., and van Berckelaer-Onnes, I. A. (2008). The central coherence account of autism revisited: Evidence from the ComFor study. *Res. Autism Spectr. Disord.* 2, 209–222. doi: 10.1016/j.rasd.2007.05.004

Nydén, A., Hagberg, B., Goussé, V., and Rastam, M. A. (2011). cognitive endophenotype of autism in families with multiple incidence. *Res. Autism Spectr. Disord.* 5, 191–200. doi: 10.1016/j.rasd.2010.03.010

O'Riordan, M. A., Plaisted, K. C., Driver, J., and Baron-Cohen, S. (2001). Superior visual search in autism. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 719–730. doi: 10.1037/0096-1523.27.3.719

Ornitz, E. M., Guthrie, D., and Farley, A. H. (1977). The early development of autistic children. *J. Autism Child Schizophr.* 7, 207–229. doi: 10.1007/BF01538999

Pellicano, E., Maybery, M., and Durkin, K. (2005). Central coherence in typically developing preschoolers: Does it cohere and does it relate to mindreading and executive control? *J. Child Psychol. Psychiatry* 46, 533–547. doi: 10.1111/j.1469-7610.2004.00380.x

Pellicano, E., Maybery, M., Durkin, K., and Maley, A. (2006). Multiple cognitive capabilities/deficits in children with an autism spectrum disorder: "Weak" central coherence and its relationship to theory of mind and executive control. *Dev. Psychopathol.* 18, 77–98. doi: 10.1017/S0954579406060056

Perkins, M. (2010). "Pragmatic impairment," in *The handbook of language and speech disorders*, eds S. Jack, Damico, Nicole Müller, J. Martin, and Ball (Chichester: Wiley-Blackwell), 227–246. doi: 10.1002/9781444318975.ch10

Pina, F., Flavia, M., and Patrizia, O. (2013). Relationship between weak central coherence and mental states understanding in children with autism and in children with ADHD. *Mediter. J. Clin. Psychol.* 1:1.

Plaisted, K., O'Riordan, M., and Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *J. Child Psychol. Psychiatry* 39:S0021963098002613. doi: 10.1017/S0021963098002613

Rapin, I., and Allen, D. (1983). Developmental language disorders: Nosologic considerations. *Neuropsychol. Lang. Read. Spell.* 55, 155–184. doi: 10.1016/B978-0-12-409680-6.50014-7

Rinehart, N. J., Bradshaw, J. L., Moss, S. A., Brereton, A. V., and Tonge, B. J. (2000). Atypical interference of local detail on global processing in high-functioning autism and asperger's disorder. *J. Child Psychol. Psychiatry* 41, 769–778. doi: 10.1111/1469-7610.00664

Rinehart, N. J., Bradshaw, J. L., Moss, S. A., Brereton, A. V., and Tonge, B. J. (2001). A deficit in shifting attention present in high-functioning autism but not Asperger's disorder. *Autism* 5, 67–80. doi: 10.1177/1362361301005001007

Santangelo, S. L., and Folstein, S. E. (1999). "Autism: A genetic perspective," in *Neurodevelopmental disorders*, ed. H. Tager-Flusberg (Cambridge, MA: MIT Press), 431–447. doi: 10.7551/mitpress/4945.003.0024

Schoemaker, M. M., van der Wees, M., Flapper, B., Verheij-Jansen, N., Scholten-Jaegers, S., and Geuze, R. H. (2001). Perceptual skills of children with developmental coordination disorder. *Hum. Mov. Sci.* 20, 111–133. doi: 10.1016/S0167-9457(01)00031-8

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Appl. Cogn. Psychol.* 16, 989–997. doi: 10.1002/acp.930

Seernani, D., Ioannou, C., Damania, K., Spindler, K., Hill, H., Foulsham, T., et al. (2020). Studying global processing in autism and attention-deficit/hyperactivity disorder with gaze movements: The example of a copying task. *PLoS One* 15:e0224186. doi: 10.1371/journal.pone.0224186

Semrud-Clikeman, M., and Hynd, G. W. (1990). Right hemisphere dysfunction in nonverbal learning disabilities: Social, academic, and adaptive functioning in adults and children. *Psychol. Bull.* 107, 196–209. doi: 10.1037/0033-2909.107.2.196

Semrud-Clikeman, M., Walkowiak, J., Wilkinson, A., and Christopher, G. (2010). Neuropsychological differences among children with asperger syndrome, nonverbal learning disabilities, attention deficit disorder, and controls. *Dev. Neuropsychol.* 35, 582–600. doi: 10.1080/87565641.2010.494747

Shah, A., and Frith, U. (1983). An islet of ability in autistic children: A research note. *J. Child Psychol. Psychiatry* 24, 613–620. doi: 10.1111/j.1469-7610.1983.tb00137.x

Syriopoulou Delli, C. K., Varveris, A., and Geronta, A. (2016). Application of the theory of mind, theory of executive functions and weak central coherence theory to individuals with ASD. *J. Educ. Dev. Psychol.* 7:102. doi: 10.5539/jedp.v7n1p102

Tager-Flusberg, H., and Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 358, 303–314. doi: 10.1098/rstb.2002.1198

Velarde, M., Echevarría, J. A. V., and Gómez Velarde, M. (2017). Trastorno de la comunicación social (pragmático) (TCS f80.82). *Pediátr. Panamá* 46, 99–104.

Vermeulen, P. (2015). Context blindness in autism spectrum disorder. *Focus Autism Other Dev. Disabl.* 30, 182–192. doi: 10.1177/1088357614528799

Viñuela, F. (2007). "Trastornos de las funciones visuoespacial y constructiva," in *Neurología de la conducta y neuropsicología*, ed. Peña-Casanova (Madrid: Médica Panamericana), 233–242.

Wang, A. T., Dapretto, M., Hariri, A. R., Sigman, M., and Bookheimer, S. Y. (2004). Neural correlates of facial affect processing in children and adolescents with autism spectrum disorder. *J. Am. Acad. Child Adolesc. Psychiatry* 43, 481–490. doi: 10.1097/00004583-200404000-00015

Wang, L., Mottron, L., Peng, D., Berthiaume, C., and Dawson, M. (2007). Local bias and local-to-global interference without global deficit: A robust finding in autism under various conditions of attention, exposure time, and visual angle. *Cogn. Neuropsychol.* 24, 550–574. doi: 10.1080/13546800701417096

Worling, D. E., Humphries, T., and Tannock, R. (1999). Spatial and emotional aspects of language inferencing in nonverbal learning disabilities. *Brain Lang.* 70, 220–239. doi: 10.1006/brln.1999.2156

Zhang, D., and Adipat, B. (2005). Challenges, methodologies, and issues in the usability testing of mobile applications. *Int. J. Hum. Comput. Interact.* 18, 293–308. doi: 10.1207/s15327590ijhc1803_3

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 24 July 2024
DOI 10.3389/fpsyg.2024.1373898

OPEN ACCESS

EDITED BY
Ann Dowker,
University of Oxford, United Kingdom

REVIEWED BY
Ada Cigala,
University of Parma, Italy
Elisabetta Conte,
University of Milano-Bicocca, Italy
Ebtesam Rababah,
Yarmouk University, Jordan

*CORRESPONDENCE
Esra Ünlüer
esra.unluer@kocaeli.edu.tr

RECEIVED 22 January 2024
ACCEPTED 03 July 2024
PUBLISHED 24 July 2024

CITATION
Ünlüer E (2024) Theory of mind skills and peer relationships in children's adjustment to preschool.
Front. Psychol. 15:1373898.
doi: 10.3389/fpsyg.2024.1373898

Theory of mind skills and peer relationships in children's adjustment to preschool

Esra Ünlüer*

Department of Preschool Education, Faculty of Education, Kocaeli University, Kocaeli, Türkiye

School adjustment affects children's future lives in many ways. This study examined the relationship between ToM skills, peer relationships, and school adjustment. Specifically, this study determined whether preschool children's school adjustment could be significantly predicted by theory of mind (ToM) skills and peer relationships. A total of 164 children aged 4 (34.5%), and 5 (38%) years of preschool attendance participated in the study. According to the research, children's age, theory of mind, peer relations, and school adjustment are closely related. It was also found that the theory of mind significantly predicted school adjustment (school liking/avoidance) and that prosocial and aggressive behavior predicted school liking.

KEYWORDS
theory of mind, peer relationship, school adjustment, school liking, preschool children

Introduction

In the preschool period, which is an important stage in children's development, children begin the process of adapting to school. At this stage, children learn to adapt to the skills, rules, and social norms necessary to establish relationships with their teachers and peers, share play materials, and follow school routines. Preschool is a developmentally appropriate period for learning the social skills that children will need for social adaptation in later years (Johnson et al., 2000). Children's positive experiences in preschool environments can form a foundation that will enable them to establish social relationships with others; however, if they fail to overcome the difficulties experienced in this period, this can be seen as the cause of their adaptation problems later in life (Ladd et al., 1999).

Development of theory of mind in early childhood

Theory of mind (ToM) is based on the child's ability to predict and understand that the behavior of others is meaningful and intentional (Rowe et al., 2001). Children's acquisition of theory of mind is a developmental process that enables them to have an increasingly sophisticated understanding of mental states (Wellman et al., 2001; Wellman and Liu, 2004). The process of acquiring the skills of the Theory of Mind begins in infancy (Repacholi and Gopnik, 1997). With the development of language skills in the second year of life, babies begin to talk about, interpret, and make sense of their own and others' emotions and experiences (Grazzani-Gavazzi, 2004).

According to Wellman and Liu (2004), the theory of mind development begins with developing an understanding of desires and then developing an understanding of beliefs. At this point, understanding belief is divided into two parts: understanding different beliefs and understanding false beliefs. From the age of three, children begin to understand that others may have different desires and that two people with different desires may feel different emotions in the same situation (Denham and Couchoud, 1990). Around the age of 4 years, the understanding of false beliefs begins to develop, and this development continues throughout childhood (Dore et al., 2018). Although rankings of theory of mind abilities that develop with age are universal, there are individual and cultural differences (Ertugrul Yasar, 2016).

There are two basic perspectives that attempt to explain the development of the theory of mind. One is the nativist perspective, which suggests that the theory of mind is largely innate. Another is the constructivist perspective (Westra and Carruthers, 2018). According to this perspective, children build their concepts of mind with the knowledge they gain about the minds of others, and as they acquire more and more data that their current concepts cannot explain, they may revise their concepts accordingly (Gopnik and Wellman, 2012, Wellman, 2014; Lane and Bowman, 2021). In line with this perspective, there are strong relationships between the knowledge children acquire from social experiences and their theory of mind (Lane and Bowman, 2021). The constructivist view is supported by evidence that the development of the theory of mind is influenced by social experience and language (Westra and Carruthers, 2018).

Recent functional imaging studies have strongly demonstrated that Theory of Mind (ToM) functions are facilitated by a highly organized and well-localized system in the brain (Drubach, 2008). Frith and Frith (2003) noted that social cognition processes are concentrated in specific regions of the brain, and structural or functional impairments in these regions can lead to significant deficits in ToM abilities. Brothers (1990) proposed a model for social cognition (encompassing ToM abilities) that consists of a three-node circuit connecting the orbitofrontal cortex, superior temporal sulcus, and amygdala. According to Brothers, if this circuit is disrupted at any point, it can result in autism, a disorder associated with significant deficits in ToM abilities.

Peer relationships in early childhood

The concept of peer relationships involves those at a similar level of development, maturity, or age; it is expressed as a set of ongoing and reciprocal relationships between individuals with compatible values, lifestyles, social relationships, and backgrounds (Gulay, 2010). Children's relationships with their peers change with age, with an increase in the amount of time spent with peers and the influence of play (Gulay, 2008). Peer interactions increase around the age of three, and as peer relationships increase around the age of four, children begin to experience rejection or acceptance (Oral, 2015). New communication skills acquired through peer interaction ensure the continuation of relationships and bring harmony and closeness. With their improved communication skills, they also produce behavioural solutions to the problems they experience with their peers (Hay, 2006). These relationships also impact social acceptance and competence in the years to be come (Walker, 2004).

Factors that influence competence in peer relationships have been studied for many years. The roots of individual differences in peer relationships lie in many variables such as family, especially parent-child relationships, and cognitive, social-emotional and social-cognitive development (Hazen and Brownell, 1999). Studies conducted with peer relationships support this. Studies have revealed the positive effects of parent-child relationship (McHale et al., 1999), cognitive development (Meece and Mize, 2009; Gulay et al., 2013), and social-emotional development (Gulay and Onder, 2013; Sen and Ozbey, 2017) on peer relationships.

Association between theory of mind and peer relation

The theory of mind, one of the social cognitive abilities that affect peer relationships, assumes that individuals are aware of both their own cognitive state and the cognitive states of other people. The theory, which focuses on mental processes in the early years of life, is also extremely important for social development because of the impact of mental processes on social relationships (Gulay et al., 2013). Many studies conducted in recent years have focused on the relationship between children's peer relationships and the theory of mind. These studies show that the theory of mind is positively related to prosocial behavior and increases peer acceptance (Caputi et al., 2012), whereas it is negatively related to negative behavior and leads to peer rejection (Banerjee and Watling, 2005; Suway et al., 2012).

Lane and Bowman (2021) suggest that when the role of social context in the development of theory of mind is taken into account, children who have more social interaction will develop their theory of mind skills more quickly because they have the opportunity to better understand the internal states of others. Conversely, less social interaction will lead to delayed theory of mind.

Theory of mind and peer relationships in school adjustment

School adjustment is highly influential on children's future lives in many ways. School adjustment predicts a range of positive outcomes for children, including future academic success and healthy socioemotional adjustment (Entwisle and Hayduk, 1988; Ramey et al., 1998; Ladd et al., 2000; Valeski and Stipek, 2001; Herndon et al., 2013; Blair and Raver, 2015). Considering the prevalence of problems related to school adaptation and their costs for the individual and society, cognitive and social competence are important interactional processes that should be examined in school adaptation studies. In particular, it is of interest for both developmental theory and educational policy to determine the factors that predict early school adjustment and children's adaptation to new school environments (e.g., primary school entry) (Ladd, 2017).

Relationships with teachers, parents, and classmates have important effects on children's ability to cope with many difficulties brought about by school adaptation, successfully adapt to this new environment, and like school (Ladd and Price, 1987; Baker, 2006; Ladd, 2017; Liew et al., 2019; Cheung et al., 2022). Children's relationships with their peers in the classroom can be seen as an important support in the difficulties young children face in adapting to school (Coie et al., 1990; Wentzel, 1999; Betts et al., 2012; Ladd, 2017). Peers make up a large part of students' social lives in the classroom and school (León and Liew, 2017), and forming close bonds with classmates can provide kindergarten students with the sense of security they need to explore and cope with the new environment (Johnson et al., 2000). There is strong evidence that children who establish high-quality relationships with their peers and are socially competent will actively participate in classroom activities, be academically successful, and adapt to school (Ladd et al., 1999, 2006; Caprara et al., 2000; Hay et al., 2004; Walker, 2004; Prinstein et al., 2005; Wood, 2007; Valiente et al., 2008; Nakamoto and Schwartz, 2010; Eggum-Wilkens et al., 2014; Hernández et al., 2016; Garner et al., 2020). Children who can establish good relationships with their peers have higher readiness for primary school (Polat and Atis-Akyol, 2016). Conversely, socially withdrawn children are often the target of peer rejection in the preschool period, which continues in primary school (Rubin et al., 1989). In addition, children who are not accepted do not participate in classroom activities, avoid school, and experience difficulties in academic areas (Buhs and Ladd, 2001; Wentzel, 2005; Klima and Repetti, 2008).

Early school adjustment involves a collection of academic, social, and regulatory skills (Pianta et al., 1997). Theory of mind, the ability to reason about the internal states of others, is an important and complex skill that influences skills in many areas of school adjustment (Brock et al., 2018). The theory of mind has significant relationships with academic (Blair and Razza, 2007; Astington and Pelletier, 2013), social (Watson et al., 1999; Capage and Watson, 2001), and regulatory skills (Aydin and Karakelle, 2016; Risnawati et al., 2023). The significant relationship between the Theory of Mind and the skills required for school adjustment may provide evidence that the Theory of Mind predicts school adjustment.

The present study

As noted above, many studies have examined the theory of mind, peer relations, and school adjustment and their relationships with each other. To the best of our knowledge, no study has simultaneously addressed and clearly defined the role of children's theory of mind and peer relations on school adjustment. The current study will contribute to our understanding of how young children's ability to understand their internal mental states and their relationships with peers affect their adjustment to school. The ages of the children in our sample (4–5 years old), peer relations, and theory of mind guide our decision to test a model in which performance is expected to have a direct impact on school adjustment. First, consistent with the existing literature, we examined associations between children's age, theory of mind, peer relationships, and school adjustment, even after controlling for common demographic characteristics. The age variable was included in the model to be evaluated. The second aim was to examine how the theory of mind and peer relations would affect school adjustment. We predicted that both theory of mind and peer relations would have direct effects on school adjustment.

Method

Participants

Participants in this study were selected using criterion sampling, a purposive sampling technique. Eligibility criteria included the absence of a) developmental delay, b) autism spectrum disorder, and c) sensory impairment. Developmental assessments were carried out in September by school counsellors using the Ministry of Education's developmental assessment form. Accordingly, the study included 164 typically developing children aged 4 and 5 years. The children participating in the study were 78 (47.5%) 4 and 86 (52.4%) 5 years old. The children who participated in the study were 85 girls and 79 boys. In the survey conducted during the first four months of the 2023–2024 school year, children aged 4-5 have been attending school for at least two months. These children attend three different public kindergartens with similar socioeconomic levels. A total of 22 preschool teachers from these three different schools participated: 10 from the 4-year-old class, and 12 from the 5-year-old class. The teachers participating in the study had at least five years of teaching experience. While 36 (46.1%) of the 4-year-old children and 76 (88.3%) of the 5-year-old children had at least 1 year of school experience.

Procedure

The start of the study was the approval of the research procedure by the Social Sciences and Humanities Ethics Committee of Kocaeli University. State kindergartens in the central district of Kocaeli were contacted to explain the aims of the study, the process, and the criteria required for the children participating in the study. The children who would participate in the study were identified through interviews with school principals and teachers. On the basis of these interviews, children were selected who had been approved by their teachers and who did not have a diagnosis (developmental delay, autism, etc.) in their individual files. The families and teachers of the children were given a participation form explaining the purpose and process of the study, and it was stated that a report about the study would be given at the end of the study.

The researcher performed the measurements in October, November, and December. The school visits started with meeting the 4–5 years old children in the last week of October, and the measurements continued throughout October, November. Completed by the teacher at the end of December. First, the School Liking and Avoidance Questionnaire was administered, starting with five sample items of the scale, and children who could not answer these items were not included in the study. Second, the Theory of Mind Battery was presented. Measurements, which took approximately 25 min, were performed individually in the guidance service's room, and the children who volunteered to participate in the study were informed about the procedure again and reminded that they could terminate the study at any time. The Profile of Peer Relations (PPR) questionnaire was introduced to the classroom teacher and filled in by the teacher for each child.

Measures

The School Liking and Avoidance Questionnaire, Theory of Mind Battery, and Profile of Peer Relations (PPR) Questionnaire were used as data collection tools.

School liking and avoidance questionnaire

The scales developed by Ladd and Price (1987) and Ladd (2017) to determine children's attitudes toward school were adapted for Turkish children by Nur and Arnas (2019). The SLAQ, which has 14 items, has a two-factor structure and evaluates school adjustment based on children's perceptions. There are five sample items at the beginning of the questionnaire; these questions are not scored. The sub-dimension of school liking (items 1, 2, 4, 6, 7, 8, 10, 11, 12) includes children's positive perceptions and feelings about school (items 2, 6, and 12 are calculated in reverse; Is school fun? Do you like being at school?) and the school avoidance sub-dimension (items 3, 5, 9, 13, 14) assesses children's desire to avoid school (Do you want to stay at home instead of coming to school?). During individual interviews with the children, they were asked to rate the items on a three-point scale ("yes," "sometimes," or "no"), which were scored as 3, 2, and 1, respectively. Total score calculated by averaging the scores for each sub-dimension. High scores for the school liking subscale indicate positive feelings about school, whereas high scores for the avoidance subscale indicate a higher desire to avoid school. Cronbach's alpha coefficient for the school liking sub-dimension of the SLAQ was determined to be.92 and.87. For the school avoidance sub-dimension (Nur and Arnas, 2019). In this study, two factors showed high internal consistency (Cronbach's α = 0.88 and 0.81).

Profile of peer relations (PPRs)

This inventory, developed by Walker et al. (2000), is used to assess each preschool child's typical social and play behaviors with their peers (Walker et al., 2002; Walker, 2005). In the inventory, teachers were asked to rate children's competence in identified social and play behaviors. The PPR, which has 25 items, included (a) items assessing the frequency of positive and negative play behaviors, such as cooperative play, verbal aggression, and physical aggression; (b) items detailing various strategies children may use when trying to participate in other children's play; and (c) items related to the rate of engagement in conflict and conflict resolution strategies. Teachers were asked to rate the extent to which children exhibited certain behaviors on a 4-point Likert-type scale ranging from 1 (rarely) to 4 (almost always). Walker et al. (2000) identified three factors in the measure: aggressive or destructive behavior, prosocial behavior, and shy or withdrawn behavior. They reported that the three factors showed high internal consistency (Cronbach's α = 0.91, 0.86, and 0.59). In this study, three factors showed high internal consistency (Cronbach's α = 0.81, 0.79, and 0.61). The questionnaire is self-explanatory, dealing with behaviors that are easily visible within preschool environments. No prior training is required by qualified early childhood teachers to complete this inventory.

Theory of mind tests

In this study, frequently used ToM tests with varying degrees of difficulty, as described by McAlister and Peterson (2006), were used. Three different types of ToM tests were used. These are false beliefs, appearance-reality, and pretend representation.

False Belief (FB): (a) the 'Sally Ann' task of Baron-Cohen et al. (1985) was used as a standardized invisible substitution task. This task was administered and scored faithfully to the original, except that the Turkish girls' names Sare and Ayse were used instead of Sally and Ann. The trials began with the introduction of the dolls' names and the materials to be used. Then, Sare put the marble in a closed basket and left the environment. In the first trial, the doll named Ayşe put the marble in a closed box, and in the second trial, she put it in the tester's pocket. Sare returned asking for her marble. The child was asked the question "Where will Sare look for her marble first?" The answer "She will look for it in the closed box" for two trails is 1 point. She scored 1 point, with a maximum of 2. (b) A standardized deceptive container task was used, the test developed by Gopnik and Astington (1988). In this task, the child was shown a closed band-aid box and asked what was inside. Then, the child was shown the pencils inside the band-aid box, and the box was closed again. Regarding the belief of others, the child was asked, [classmate's name] will come next. He/she does not know what is in this box. When I show him/her the box, what will he/she tell me about what is in the box? In relation to his/her own false belief, he/she was asked, "When I first showed this box to you—before you saw the contents of the box—what was in the box?'. For the two questions, the answer "bandaid" is worth 1 point. The maximum score for this task is 2. The sum of the scores obtained from these two tasks constitutes the total false belief (TFB) score. The total can vary between 0 and 4. Two false belief scores were found to be $r = 0.50$, $p < 0.001$ in the McAlister and Peterson (2006) study and $r = 0.48$, $p < 0.001$in this study.

Pretend Representation (PR): The child was shown a real potato. The child was asked to pretend that the potato was soap. After a short period of role play, the game was finished and the child was asked: "What is this really?' (1 point for potato answer, 0 point for soap answer) and "What did we pretend it was?" (1 point for soap answer, 0 point for potato answer). After the child answered, a real soap and an irrelevant object (tennis ball) were placed on the table. "Which one did we pretend to be soap?" (1 point if it shows potato, 0 points if it shows soap or other objects) and "Which one is really soap?" (1 if he/she shows soap, 0 if he/she shows potato or other objects). The same test was repeated with a real banana. The banana was used as a phone in the role play, and a real cell phone was used at the end of the test. The total score was halved so that the children's total score for each task was 0-4 points. The total score was halved so that the children's scores for the two tasks would be similar to that of the false belief task. Pretend representation scores were found to be $r = 0.31$, $p < 0.01$ in the McAlister and Peterson (2006) study and $r = 0.35$, $p < 0.01$ in this study.

Appearance-Reality (AR): This test was based on Flavell (1986) procedure. In the first task, unlike the original, an orange was used

TABLE 1 Children's performance on the ToM, peer relations, and school adjustment.

		Range	Age Group		Total
			4 age	5 age	
			Main Score (SD)		
Theory of Mind	False Belief	0–4	2.12 (0.93)	3 (0.92)	
	Pretend Represent.	1–4	3.3 (0.97)	3.46 (0.74)	
	Appearance reality	0–2	0.57 (0.71)	1.51 (0.5)	
	Total ToM	0–1	0.53 (0.17)	0.79 (0.16)	
Profile of Peer Relations	Aggressive/Disruptive	1–4	1.61 (0.41)	1.51 (0.31)	
	Prosocial Behavior	1–4	2.98 (0.33)	3.42 (0.46)	
	Shy/Withdrawn	1–4	1.77 (0.42)	1.40 (0.4)	
School Adjustment	School liking	9–27	22.05 (2.34)	26 (1.16)	
	School avoidance	5–15	5.9 (1.27)	9.4 (2.12)	

instead of an apple. The children were shown an orange made of wax and asked what it was. After the children answered "orange," a candle was lit and it was shown that it was actually a candle. They were then asked "What is this really?" (reality test) and "When you look at this now, does it look like a [candle] or [an orange]?" (appearance test). The child who answered both test questions correctly scored 1 point, but the child who gave only one correct answer scored 0 points. The same trial was repeated with a ballpoint pen that actually looked like a flower. The sum of the scores from these two tasks is the total AR score. The total score can range from 0 to 2. The two appearance-reality scores were found to be $r = 0.34$, $p < 0.01$ in the McAlister and Peterson (2006) study and $r = 0.38$, $p < 0.001$ in this study.

Total ToM score

To calculate the total ToM score, first the relationships between the three tasks were examined and then the total score was calculated. The total FB scores of the children were found to be significantly correlated with AR ($r = 0.69$, $p < 0.001$) and PR ($r = 0.56$, $p < 0.001$); the total PR score was found to be significantly correlated with the total AR score ($r = 0.23$, $p < 0.001$). To create the total ToM score, first, the sums of FB and PR totals were divided by 4 and the AR total by 2 to compute the proportion score for each component. The total score was averaged to create a total ToM score that could vary between 0 and 1. In this study, Total ToM showed acceptable internal consistency (Cronbach's $\alpha = 0.76$).

Statistical analyses

All statistical analyses were performed using SPSS version 22. Descriptive analyses were performed on the scores of the children using the data collection tools, and the mean and standard deviation were calculated. To test the normality of the distribution of the data, Skewness and Kurtosis Test was performed ToM (Skewness −0.089; Kurtosis −0.675), School Liking (Skewness −0.689; Kurtosis −0.961), School Avoidance (Skewness 0.717; − Kurtosis −0.482), Aggressive/Disruptive Behaviour (Skewness −0.183; Kurtosis −0.746), Prosocial Behaviour (Skewness.307; Kurtosis −1.274), Shy/Withdrawn Behaviour (Skewness 0.358; Kurtosis −0.860) and it was determined that the distribution was between 1.5 and 1.5. Because the data were within the normal distribution range, parametric tests were applied. The Pearson Product Moment Correlation Coefficient was calculated to reveal the relationship among children's age, ToM performance, peer relations, and school adjustment. Multiple linear regression analysis was performed to determine the predictive power of ToM and peer relations on school adjustment.

Results

Descriptive analysis

The descriptive statistics of the study variables are presented in Table 1.

Correlational analyses

Table 2 provides an overview of the correlations among the variables included in the study.

Table 2 shows that children's age is significantly positively related to ToM ($p < 0.001$), prosocial behavior ($p < 0.001$) and school liking ($p < 0.001$), and is significantly negatively related to shy or withdrawn behavior ($p < 0.001$) and school avoidance ($p < 0.001$).

Regression analysis

A regression analysis was conducted to investigate whether children's peer relationships and Theory of Mind scores explain the variance in school liking and avoidance scores. The regression model was designed to evaluate the variance in school liking/avoidance scores in three successive steps. In the first step (Step 1), the role of age was assessed. Subsequently (Step 2), the impact of three dimensions of peer relationships (prosocial, aggressive, and shy behaviors) was examined by entering them into the regression model. Finally, in Step 3, the effect of Theory of Mind was evaluated.

The age of the children was entered in Step 1. The model was statistically significant, $F(1,162) = 204.67$, $p < 0.000$, explaining

TABLE 2 Correlations among children's age, ToM, peer relations, and school adjustment.

	Theory of Mind	Aggressive /Disruptive	Prosocial Behavior	Shy /Withdrawn	School liking	School avoidance
Age	0.626**	−0.108	0.462**	−0.392**	0.747**	−0.728**
Theory of Mind		−0.266**	0.784**	−0.704**	0.717**	−0.671**
Aggressive/Disruptive			−0.294**	0.034	−0.387**	0.184*
Prosocial Behavior				−0.689**	0.666**	−0.459**
Shy/Withdrawn					−0.443**	0.463**
School liking						−0.820**
School avoidance						–

*$p < 0.01$ **$p < 0.001$.

TABLE 3 Regression outcomes for target variable School Liking.

	β	SE	t	p
Step 1				
Age	0.747	0.280	14.307	0.000
Step 2				
Age	0.57	0.248	12.340	0.000
Aggressive/Disruptive	−0.219	0.350	−4.989	0.000
Prosocial Behavior	0.366	0.361	5.822	0.000
Shy/Withdrawn	−0.041	0.345	0.697	0.487
Step 3				
Age	0.516	0.279	9.911	0.000
Aggressive/Disruptive	−0.201	0.353	−4.546	0.000
Prosocial Behavior	0.296	0.402	4.224	0.000
Shy/Withdrawn	0.093	0.370	1.487	0.139
ToM	0.174	1.012	2.162	0.03

TABLE 4 Regression outcomes for target variable School Avoidance.

	β	SE	t	p
Step 1				
Age	−0.728	0.264	−13.519	0.000
Step 2				
Age	−0.639	0.286	−10.934	0.000
Aggressive/Disruptive	0.114	0.405	2.063	0.061
Prosocial Behavior	0.025	0.416	0.310	0.757
Shy/Withdrawn	0.225	0.399	3.061	0.003
Step 3				
Age	−0.512	0.310	−8.078	0.000
Aggressive/Disruptive	0.072	0.393	1.342	0.181
Prosocial Behavior	0.188	0.447	2.205	0.079
Shy/Withdrawn	0.103	0.412	1.357	0.177
ToM	−0.406	1.126	−4.148	0.000

approximately 55% of the variance. Only the age variable accounted for the variability in school like scores (β = 0.74, $p < 0.001$). In Step 2, F(4,159) = 111.07, $p < 0.000$, the inclusion of three dimensions of peer relationships resulted in a statistically significant increase in explained variance ($R_2 = 0.73$). Specifically, it was found that the dimensions of prosocial behaviors (β = 0.25, $p < 0.000$) and aggressive behaviors (β = −0.12, $p < 0.000$) of peer relationships had a significant impact on school liking scores. Finally, when Theory of Mind (ToM) was entered into the regression model, F(5,158) = 91.85, $p < 0.000$, it increased the explained variance by 1% ($R_2 = 0.74$). The beta weight for ToM is (β = 0.23, $p < 0.05$) (see Table 3). This result indicates that prosocial and aggressive behaviors, along with Theory of Mind, predict school liking scores.

The age of the children was entered together in Step 1. The model was statistically significant, F(1,162) = 182.77, $p < 0.000$, explaining approximately 53% of the variance. Only the age variable accounted for the variability in school avoidance scores (β = −0.72, $p < 0.000$). In Step 2, F(4,2159) = 54.68, p < 0.000, the inclusion of three dimensions of peer relationships resulted in a statistically significant increase in explained variance ($R_2 = 0.57$). Specifically, it was found that the dimension of peer relationships' shyness (β = 0.22, $p < 0.001$) had a significant impact on school avoidance. Finally, when Theory of Mind (ToM) was entered into the regression model, F(5,220) = 51.64, $p < 0.000$, it increased the explained variance by 5% ($R_2 = 0.62$). The beta weight for ToM (β = −0.40, $p < 0.000$) indicated that it should be considered a robust determinant of school avoidance scores (see Table 4). When accounting for the effect of Theory of Mind in the model, it can be observed that the relationship between shyness and school avoidance scores loses its statistical significance. This result demonstrates the significant role of Theory of Mind in the relationship between shyness and school avoidance.

Discussion

The main aim of this study was to investigate both theory of mind and peer relations in children's adjustment to school in early childhood. Because of the study, two main findings were obtained. First, there were significant relationships between the investigated variables even after controlling for age. Second, it was found that theory of mind significantly predicted school adjustment (school liking/avoidance) regardless of age, prosocial behaviors positively and aggressive behaviors negatively predicted school liking.

Relationships among age, theory of mind, peer relation, and school adjustment

In this study, statistically significant correlations were observed between the investigated variables. Consistent with the literature, this study shows that overall performance on ToM tasks increases with age (Gopnik and Slaughter, 1991; Wellman et al., 2001; Slaughter et al., 2002; Newton and Jenvey, 2011; Conte et al., 2019). The study found, as expected, that the Theory of Mind was strongly related to peer relationships. Consistent with previous research, there were positive relationships between the theory of mind and prosocial behavior and negative relationships between aggression and shyness. Imuta et al. (2016) reviewed 76 studies conducted with children aged 2–12 years and found a positive relationship between ToM and prosocial behaviors, similar to our study. Early understanding of the basic concepts of the theory of mind supports the development of prosocial behavior and indirectly mediates peer acceptance (Denham et al., 1990; Capage and Watson, 2001; Caputi et al., 2012). The basic concepts in the theory of mind are beliefs, desires, and intentions that are used to understand the behavior of others and predict how they will behave (Kloo et al., 2010). The theory of mind is necessary for children to learn basic concepts of prosocial behavior such as sharing and cooperation. Delays in the acquisition of these basic concepts are thought to underlie aggressive and destructive behavior in children (Capage and Watson, 2001; Wellman et al., 2011; Shakoor et al., 2012; Lane et al., 2013).

Previous studies on shy children have shown different results regarding their theory of mind abilities. In some studies, introverted children learn the contents of others' minds through social observation and have a high theory of mind (Wellman et al., 2011; LaBounty et al., 2017), whereas in others, behavioral reactions, social anxiety, and fear-related behaviors are reported. Shy children have a less developed theory of mind (Banerjee and Henderson, 2001; Suway et al., 2012, Lane et al., 2013). Shy individuals, who are cautious and anxious about the situations created by social environments, experience a conflict of approach and avoidance in such situations. While shy children want to engage in social interaction, they avoid interaction because of the fear and anxiety they experience (Coplan et al., 2004; Coplan and Armer, 2007). Avoidance of interaction causes shy children to have many problems adjusting to school and to avoid school (Wu et al., 2015).

ToM and peer relationships in school adjustment

In this study, theory of mind and prosocial behavior positively predicted school liking and aggressive behavior negatively; only theory of mind was found to negatively predict school avoidance. These findings agree with previous studies showing that theory of mind (Bolnick, 2008) and peer relationships (Gulay and Erten, 2013; Aydogdu, 2022) have important effects on children's adjustment to school. Pre-school education presents many new challenges for children, such as meeting the demands of the environment, building relationships with teachers and peers, expanding their social circle, and dealing with social relationship problems. The theory of mind facilitates school adjustment by helping to develop some of the basic skills needed to cope with these difficulties (Chi et al., 2018; Brock et al., 2019). In addition, Dunn (1995) stated that children's adjustment to school, which is a social environment, may be related to some characteristics of their understanding of others' internal states, and that children with better ability to understand others should like school more. There are studies that show that the Theory of Mind has strong relationships with other skills that affect school adjustment, apart from social relationships. The strong relationships of the Theory of Mind with cognitive skills (Lockl and Schneider, 2007; Lecce et al., 2017), such as metacognitive and language skills (Conte et al., 2019), support its effect on school adjustment.

Positive peer relationships in the preschool years support children in the school environment and help them adjust to school. This study also shows that prosocial and aggressive behaviors are effective in children's liking of school. This finding of our study is in line with the results of previous studies. In fact, Gulay and Erten (2013) and Aydogdu (2022) reported in their studies with preschool children that peer relationships predict liking for school. Children who display aggressive and disruptive behavior are often rejected by their peers (Slaughter et al., 2002; Gunnar et al., 2003) and have difficulty adjusting to school. In contrast, children who engage in positive social behavior are accepted by their peers and adapt to school more easily (Chung-Hall and Chen, 2010). In addition, this study showed that shyness did not have a direct effect on liking school. This finding was not consistent with studies showing that the combination of shyness and reserve predicts a range of social, emotional, and school adjustment difficulties in preschool (Chen et al., 1995; Coplan et al., 2008; Wu et al., 2015). Unlike previous studies examining the relationship between school adjustment and withdrawn behavior, this study includes the variable of theory of mind closely associated with withdrawn behaviors (Banerjee and Henderson, 2001). This suggests the possibility that shyness predicts school adjustment through the theory of mind.

Implications and limitations

In summary, this study found that theory of mind, peer relations, and school adjustment are closely related; that liking school, which is a sub-dimension of school adjustment, is predicted by theory of mind as well as prosocial and aggressive behavior; and that the most important predictor of young children's school adjustment is theory of mind. To the best of our knowledge, this is the first study to examine the relationships between these variables in a single model with children under the age of six. It is recognized that this study has some strengths and some limitations.

First, the cross-sectional nature of the study is a limitation in understanding the changes and effects of the variables in subsequent years. Second, the fact that the data in the study were collected very soon after the children started school and the possibility that the teachers, who are the source of information about the children's peer relationships, do not know the children well enough is another important limitation of the study. Although this study contributes to our understanding of young children's theory of mind, peer relationships, school adjustment, and their interrelationships, a longitudinal study in which measurements are repeated at different times will contribute to our understanding of long-term changes or continuity in these variables. This study attempted to determine whether there is a direct relationship between the theory of mind, peer relationships, and school adjustment. However, future studies should examine a model that incorporates additional variables potentially impacting children's adjustment to school, such as teacher-child relationships, emotion regulation and language skills. Another limitation of this study is that information on school adjustment and peer relationships was obtained from a single source. The use of multiple sources of information (peers, parents, etc.) in future studies will further contribute to the validity of the findings. Lastly, in this study, children with typical development were examined, and their developmental assessments were conducted by their teachers. However, no developmental tests were administered by the researcher. This raises the possibility that children with undiagnosed developmental delays or autism spectrum disorder might have been included in the study. Future research should apply standard developmental tests to more accurately determine the developmental status of the children included, thereby ensuring more reliable and generalizable results.

Theory of mind skills are essential for children to develop positive social relationships, to develop skills in coping with social problem situations, and to adapt to school (Guven et al., 2019). Theory-of-mind-based educational intervention programs improve children's theory of mind skills and prosocial behavior (Gozun Kahraman, 2012). In preschool classrooms, the frequent use of theory of mind concepts (belief, desire, intention, etc.) by teachers should be supported by enriched learning centers and activities (play, drama, storytelling, etc.) that allow children to use these concepts. It facilitates school adjustment by preventing negative peer relationships resulting from the lack and inadequacy of theory skills.

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 the Kocaeli University Social and Humanities Ethics Committee, Türkiye (approval number: E-94094268-204.01.07-535197). 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

EU: Writing—original draft, Writing—review and editing.

Funding

The author declares that no financial support was received for the research, authorship, and/or publication 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.

Publisher's note

References

Astington, J. W., and Pelletier, J. (2013). "Theory of mind, language, and learning in the early years: Developmental origins of school readiness," in *The development of social cognition and communication*, eds B. D. Homer and C. S. Tamis-LeMonda (London: Psychology Press), 205–230.

Aydin, M. S., and Karakelle, S. (2016). 3-4 yaş çocuklarında zihin kuramı ve kendini düzenleme becerisinin yardım davranışları üzerindeki etkisi. [The effects of theory of mind and self-regulation skills on helping behaviors in 3-4-year-old children]. *Nesne Psikol. Dergisi* 4, 205–226.

Aydogdu, F. (2022). Okul Öncesi dönem çocuklarinin akran ilişkileri ile okula uyumlari arasindaki ilişkide öz-düzenleme becerilerinin aracilik rolü [The intermediary role of self-regulation skills in the correlation between peer relations and school adaptation in preschool children]. *Eğitim Bilim* 47:11353. doi: 10.15390/EB.2022.11353

Baker, J. A. (2006). Contributions of teacher–child relationships to positive school adjustment during elementary school. *J. Sch. Psychol.* 44, 211–229. doi: 10.1016/j.jsp.2006.02.002

Banerjee, R., and Henderson, L. (2001). Social-cognitive factors in childhood social anxiety: A preliminary investigation. *Soc. Dev.* 10, 558–572. doi: 10.1111/1467-9507.00180

Banerjee, R., and Watling, D. (2005). Children's understanding of faux pas: Associations with peer relations. *Hellenic J. Psychol.* 2, 27–45. doi: 10.1111/j.1467-8624.2011.01669.x

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

Betts, L. R., Rotenberg, K. J., Trueman, M., and Stiller, J. (2012). Examining the components of children's peer liking as antecedents of school adjustment. *Br. J. Dev. Psychol.* 30, 303–325. doi: 10.1111/j.2044-835X.2011.02041.x

Blair, C., and Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Annu. Rev. Psychol.* 66, 711–731. doi: 10.1146/annurev-psych-010814-015221

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

Bolnick, R. (2008). *Contributions of children's theory of mind, inhibitory control and social competence to school adjustment*. Tempe, AZ: Arizona State University.

Brock, L. L., Kim, H., Gutshall, C. C., and Grissmer, D. W. (2018). The development of theory of mind: Predictors and moderators of improvement in kindergarten. *Early Child Dev. Care* 189:3481. doi: 10.1080/03004430.2017.1423481

Brock, L. L., Kim, H., Kelly, C. L., Mashburn, A. J., and Grissmer, D. W. (2019). Theory of mind, directly and indirectly, facilitates kindergarten adjustment via verbal ability, executive function, and emotion knowledge. *Psychol. Sch.* 56, 176–193. doi: 10.1002/pits.22216

Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts Neurosci.* 1, 27–51. doi: 10.7551/mitpress/3077.003.0029

Buhs, E. S., and Ladd, G. W. (2001). Peer rejection as antecedent of young children's school adjustment: An examination of mediating processes. *Dev. Psychol.* 37:550. doi: 10.1037/0012-1649.37.4.550

Capage, L., and Watson, A. C. (2001). Individual differences in theory of mind, aggressive behavior, and social skills in young children. *Early Educ. Dev.* 12, 613–628. doi: 10.1207/s15566935eed1204_7

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., and Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychol. Sci.* 11, 302–306. doi: 10.1111/1467-9280.00260

Caputi, M., Lecce, S., Pagnin, A., and Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: The role of prosocial behavior. *Dev. Psychol.* 48:257. doi: 10.1037/a0025402

Chen, X., Rubin, K. H., and Li, B. (1995). Social and school adjustment of shy and aggressive children in China. *Dev. Psychopathol.* 7, 337–349. doi: 10.1017/S0954579400006544

Cheung, S. K., Cheng, W. Y., Cheung, R. Y., Lau, E. Y. H., and Chung, K. K. H. (2022). Home learning activities and parental autonomy support as predictors of pre-academic skills: The mediating role of young children's school liking. *Learn. Individ. Differ.* 94:102127. doi: 10.1016/j.lindif.2022.102127

Chi, S. A., Kim, S., and Kim, N. H. (2018). A study of school adjustment related variables of young children. *South Afr. J. Educ.* 38, 1–9.

Chung-Hall, J., and Chen, X. (2010). Aggressive and prosocial peer group functioning: Effects on children's social, school, and psychological adjustment. *Soc. Dev.* 19, 659–680. doi: 10.1111/j.1467-9507.2009.00556.x

Coie, J. D., Dodge, K. A., and Kupersmidt, J. B. (1990). "Peergroup behavior and social status," in *Peer rejection in childhood*, eds S. R. Asher and J. D. Coie (New York, NY: Cambridge University Press), 17–59.

Conte, E., Ornaghi, V., Grazzani, I., Pepe, A., and Cavioni, V. (2019). Emotion knowledge, theory of mind, and language in young children: Testing a comprehensive conceptual model. *Front. Psychol.* 10:475477. doi: 10.3389/fpsyg.2019.02144

Coplan, R. J., and Armer, M. (2007). A "multitude" of solitude: A closer look at social withdrawal and nonsocial play in early childhood. *Child Dev. Perspect.* 1, 26–32. doi: 10.1111/j.1750-8606.2007.00006.x

Coplan, R. J., Arbeau, K. A., and Armer, M. (2008). Don't fret, be supportive! Maternal characteristics linking child shyness to psychosocial and school adjustment in kindergarten. *J. Abnorm. Child Psychol.* 36, 359–371. doi: 10.1007/s10802-007-9183-7

Coplan, R. J., Prakash, K., O'Neil, K., and Armer, M. (2004). Do you "want" to play? Distinguishing between conflicted shyness and social disinterest in early childhood. *Dev. Psychol.* 40, 244–258. doi: 10.1037/0012-1649.40.2.244

Denham, S. A., and Couchoud, E. A. (1990). Young preschoolers' understanding of emotions. *Child Study J.* 20, 171–192.

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

Dore, R. A., Amendum, S. J., Golinkoff, R. M., and Hirsh-Pasek, K. (2018). Theory of mind: A hidden factor in reading comprehension? *Educ. Psychol. Rev.* 30, 1067–1089. doi: 10.1007/s10648-018-9443-9

Drubach, D. A. (2008). The purpose and neurobiology of theory of mind functions. *J. Religion Health* 47, 354–365. doi: 10.1007/s10943-007-9155-9

Dunn, D. E. (1995). *The effect of structured task presentations and reinforcement on attention, achievement, and attitude of selected high school choirs*. Baton Rouge, LA: Louisiana State University and Agricultural & Mechanical College.

Eggum-Wilkens, N. D., Valiente, C., Swanson, J., and Lemery-Chalfant, K. (2014). Children's shyness, popularity, school liking, cooperative participation, and internalizing problems in the early school years. *Early Childh. Res. Q.* 29, 85–94. doi: 10.1016/j.ecresq.2013.10.002

Entwisle, D. R., and Hayduk, L. A. (1988). Lasting effects of elementary school. *Sociol. Educ.* 36, 147–159. doi: 10.2307/2112624

Ertugrul Yasar, Z. (2016). *Zihin kurami gelişim sürecinin incelenmesi [Investigation of theory of mind development]*. Ph.D. thesis. İstanbul: İstanbul Üniversitesi Sosyal Bilimler Enstitüsü.

Flavell, J. H. (1986). The development of children's knowledge about the appearance-reality distinction. *Am. Psychol.* 41, 418–425. doi: 10.1037//0003-066x.41.4.418

Frith, U., and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. Lond.* 358, 459–473.

Garner, P. W., Dunsmore, J. C., and Bassett, H. H. (2020). Direct and indirect pathways to early school adjustment: Roles of young children's mental representations and peer victimization. *Early Childh. Res. Q.* 51, 100–109. doi: 10.1016/j.ecresq.2019.09.001

Gopnik, A., and Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Dev.* 59, 26–37. doi: 10.1111/j.1467-8624.1988.tb03192.x

Gopnik, A., and Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Dev.* 62, 98–110. doi: 10.1111/j.1467-8624.1991.tb01517.x

Gopnik, A., and Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychol. Bull.* 138, 1085–1108. doi: 10.1037/a0028044

Gozun Kahraman, O. (2012). *Zihin kuramına dayalı eğitim programının 48-60 aylık çocukların bilişsel bakış açısı becerileri ve prososyal davranışları üzerindeki etkisi [Investigation of the theory of mind based training program's effects on the 48-60 months-old children's cognitive perspective taking skills and prosocial behaviors]*. Ph.D. thesis. Ankara: Gazi Üniversitesi Eğitim Bilimleri Enstitüsü.

Grazzani-Gavazzi, I. (2004). *La competenza emotiva. Studi e ricerche nel ciclo di vita*. Milano: Unicopli.

Gulay, H. (2008). *5-6 yaş çocuklarına yönelik akran ilişkileri ölçeklerinin geçerlik güvenirlik çalışmaları ve akran ilişkilerinin çeşitli değişkenler açısından incelenmesi [Standatization of a scale for measuring peer relations among 5-6 years old children and studying the relations between some familial variables and peer relations of children at this age]*. Ph.D. thesis. İstanbul: Marmara Universitesi Eğitim Bilimleri Enstitüsü.

Gulay, H. (2010). *Okul Öncesi Dönemde Akran İlişkileri, [Peer relationships in preschool]*. Ankara: Pegem Akademi.

Gulay, H. O., Önder, A., Seçer, Z., and Erten, H. (2013). Anne tutumlarının 5-6 yaş çocuklarının sosyal becerilerini ve okula uyumlarının yordayıcı etkisi. *Selçuk Ünivers. Sosyal Bilimler Enst. Dergisi* 29, 143–152.

Gulay, H., and Erten, H. (2013). 5-6 yaş çocuklarinin akran ilişkileri ve sosyal konumlarinin okula uyum düzeyleri üzerindeki yordayici etkisi (boylamsal çalışma). *Selçuk Ünivers. Sos. Bilimler Enst. Dergisi* 30, 153–163.

Gulay, H., and Onder, A. (2013). A study of social–emotional adjustment levels of preschool children in relation to peer relationships. *Education* 41, 514–522. doi: 10.1080/03004279.2011.609827

Gunnar, M. R., Sebanc, A. M., Tout, K., Donzella, B., and Van Dulmen, M. M. (2003). Peer rejection, temperament, and cortisol activity in preschoolers. *Dev. Psychobiol.* 43, 346–368. doi: 10.1002/dev.10144

Guven, Y., Ayvaz, E., and Goktas, I. (2019). Okul öncesi dönem çocuklarının zihin kuramı ve sosyal problem çözme becerileri arasındaki ilişkinin incelenmesi [Investigation of the relationship between theory of mind and socialproblem solving skills of preschool children]. *Erken Çocukluk Çalışmaları Dergisi* 3, 76–97. doi: 10.24130/eccd-jecs.1967201931130

Hay, D. F. (2006). Yours and mine: Toddlers' talk about possessions with familiar peers. *Br. J. Dev. Psychol.* 24, 39–52. doi: 10.1348/026151005X68880

Hay, D. F., Payne, A., and Chadwick, A. (2004). Peer relations in childhood. *J. Child Psychol. Psychiatry* 45, 84–108. doi: 10.1046/j.0021-9630.2003.00308.x

Hazen, N., and Brownell, C. (1999). Peer relationships in early childhood: Currents trends and future direction. *Early Educ. Dev.* 10, 233–240. doi: 10.1207/s15566935eed1003_1

Hernández, M. M., Eisenberg, N., Valiente, C., VanSchyndel, S. K., Spinrad, T. L., Silva, K. M., et al. (2016). Emotional expression in school context, social relationships, and academic adjustment in kindergarten. *Emotion* 16:553. doi: 10.1037/emo0000147

Herndon, K. J., Bailey, C. S., Shewark, E. A., Denham, S. A., and Bassett, H. H. (2013). Preschoolers' emotion expression and regulation: Relations with school adjustment. *J. Genet. Psychol.* 174, 642–663. doi: 10.1080/00221325.2012.759525

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. doi: 10.1037/dev0000140

Johnson, C., Ironsmith, M., Snow, C. W., and Poteat, G. M. (2000). Peer acceptance and social adjustment in preschool and kindergarten. *Early Childh. Educ. J.* 27, 207–212. doi: 10.1023/B:ECEJ.0000003356.30481.7a

Klima, T., and Repetti, R. L. (2008). Children's peer relations and their psychological adjustment: Differences between close friendships and the larger peer group. *Merrill Palmer Q.* 8, 151–178. doi: 10.1353/mpq.2008.0016

Kloo, D., Perner, J., and Gritzer, T. (2010). "Object-based set-shifting in preschoolers: Relations to theory of mind," in *Self and social regulation: Social interaction and the development of social understanding and executive functions*, eds B. W. Sokol, U. Müller, J. I. M. Carpendale, A. R. Young, and G. Iarocci (New York, NY: Oxford University Press), 193–217.

LaBounty, J., Bosse, L., Savicki, S., King, J., and Eisenstat, S. (2017). Relationship between social cognition and temperament in preschool-aged children. *Infant Child Dev.* 26:e1981. doi: 10.1002/icd.1981

Ladd, G. W. (2017). Having friends, keeping friends, making friends, and being liked by peers in the classroom: Predictors of children's early school adjustment? *Interpers. Dev.* 61, 203–222.

Ladd, G. W., and Price, J. M. (1987). Predicting children's social and school adjustment following the transition from preschool to kindergarten. *Child Dev.* 251, 1168–1189. doi: 10.2307/1130613

Ladd, G. W., Birch, S. H., and Buhs, E. S. (1999). Children's social and scholastic lives in kindergarten: Related spheres of influence? *Child Dev.* 70, 1373–1400. doi: 10.1111/1467-8624.00101

Ladd, G. W., Buhs, E. S., and Seid, M. (2000). Children's initial sentiments about kindergarten: Is school liking an antecedent of early classroom participation and achievement? *Merrill Palmer Q.* 28, 255–279.

Ladd, G. W., Herald, S. L., and Kochel, K. P. (2006). School readiness: Are there social prerequisites? *Early Educ. Dev.* 17, 115–150. doi: 10.1207/s15566935eed1701_6

Lane, J. D., and Bowman, L. C. (2021). How children's social tendencies can shape their theory of mind development: Access and attention to social information. *Dev. Rev.* 61:100977. doi: 10.1016/j.dr.2021.100977

Lane, J. D., Wellman, H. M., Olson, S. L., Miller, A. L., Wang, L., and Tardif, T. (2013). Relations between temperament and theory of mind development in the United States and China: Biological and behavioral correlates of preschoolers' false-belief understanding. *Dev. Psychol.* 49:825. doi: 10.1037/a0028825

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

León, J., and Liew, J. (2017). Profiles of adolescents' peer and teacher relatedness: Differences in well-being and academic achievement across latent groups. *Learn. Individ. Differ.* 54, 41–50. doi: 10.1016/j.lindif.2017.01.009

Liew, J., Valiente, C., Hernández, M. M., and Abrera, D. (2019). "Self-regulation and reactivity, school-based relationships, and school engagement and achievement," in *The SAGE handbook of developmental psychology and early childhood education*, eds D. Whitebread, V. Grau, K. Kumpulainen, M. M. McClelland, N. E. Perry, and D. Pino-Pasternak (Thousand Oaks, CA: SAGE Publishing), 42–62.

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

McAlister, A., and Peterson, C. C. (2006). Mental playmates: Siblings, executive functioning and theory of mind. *Br. J. Dev. Psychol.* 24, 733–751. doi: 10.1348/026151005X70094

McHale, J. P., Johnson, D., and Sinclair, R. (1999). Family dynamics, preschoolers' family representations, and preschool peer relationships. *Early Educ. Dev.* 10, 373–401. doi: 10.1207/s15566935eed1003_8

Meece, D., and Mize, J. (2009). Cognitive representations of peer relationships: Linkages with discrete social cognition and social behavior. *Early Child Dev. Care* 179, 539–557. doi: 10.1080/13619460701269609

Nakamoto, J., and Schwartz, D. (2010). Is peer victimization associated with academic achievement? A meta-analytic review. *Soc. Dev.* 19, 221–242. doi: 10.1111/j.1467-9507.2009.00539.x

Newton, E., and Jenvey, V. (2011). Play and theory of mind: Associations with social competence in young children. *Early Child Dev. Care* 181, 761–773. doi: 10.1080/03004430.2010.486898

Nur, I., and Arnas, Y. A. (2019). Psychometric properties of the school liking and avoidance questionnaire in a turkish preschool sample. *Int. J. Assess. Tools Educ.* 6, 539–554. doi: 10.21449/ijate.571664

Oral, T. (2015). "Kişilik gelişimi [Personality development]," in *Erken çocukluk döneminde gelişim [Development in early childhood]*, ed. E. Arslan (Ankara: Eğiten Kitap Yayıncılık), 85–109.

Pianta, R. C., Nimetz, S. L., and Bennett, E. (1997). Mother-child relationships, teacher-child relationships, and school outcomes in preschool and kindergarten. *Early Childh. Res. Q.* 12, 263–280. doi: 10.1016/S0885-2006(97)90003-X

Polat, O., and Atis-Akyol, N. (2016). Analyzing the relationship between school readiness and peer relations of five-year-old children. *Int. J. New Trends Educ. Impl.* 7, 29–39.

Prinstein, M. J., Cheah, C. S., and Guyer, A. E. (2005). Peer victimization, cue interpretation, and internalizing symptoms: Preliminary concurrent and longitudinal findings for children and adolescents. *J. Clin. Child Adolesc. Psychol.* 34, 11–24. doi: 10.1207/s15374424jccp3401_2

Ramey, S. L., Lanzi, R. G., Phillips, M. M., and Ramey, C. T. (1998). Perspectives of former Head Start children and their parents on school and the transition to school. *Element. Sch. J.* 98, 311–327. doi: 10.1086/461898

Repacholi, B. M., and Gopnik, A. (1997). Early reasoning about desires: Evidence from 14-and 18-month-olds. *Dev. Psychol.* 33:12. doi: 10.1037/0012-1649.33.1.12

Risnawati, E., Wardani, L. M. I., Saputra, A. H., and Pramitasari, M. (2023). Theory of mind, roles, and the development of emotion regulation in early childhood. *Jurnal Pendidikan Usia Dini* 17, 194–204. doi: 10.21009/JPUD.172.01

Rowe, A. D., Bullock, P. R., Polkey, C. E., and Morris, R. G. (2001). Theory of mind'impairments and their relationship to executive functioning following frontal lobe excisions. *Brain* 124, 600–616. doi: 10.1093/brain/124.3.600

Rubin, K. H., Hymel, S., and Mills, R. S. (1989). Sociability and social withdrawal in childhood: Stability and outcomes. *J. Pers.* 57, 237–255. doi: 10.1111/j.1467-6494.1989.tb00482.x

Sen, B., and Ozbey, S. (2017). Investigation of the relationship between emotional intelligence levels and actual relations in pre-school children. *Educ. Sci.* 12, 40–57. doi: 10.12739/NWSA.2017.12.1.1C0668

Shakoor, S., Jaffee, S. R., Bowes, L., Ouellet-Morin, I., Andreou, P., Happé, F., et al. (2012). A prospective longitudinal study of children's theory of mind and adolescent involvement in bullying. *J. Child Psychol. Psychiatry* 53, 254–261. doi: 10.1111/j.1469-7610.2011.02488.x

Slaughter, V., Dennis, M. J., and Pritchard, M. (2002). Theory of mind and peer acceptance in preschool children. *Br. J. Dev. Psychol.* 20, 545–564. doi: 10.1348/026151002760390945

Suway, J. G., Degnan, K. A., Sussman, A. L., and Fox, N. A. (2012). The relations among theory of mind, behavioral inhibition, and peer interactions in early childhood. *Soc. Dev.* 21, 331–342. doi: 10.1111/j.1467-9507.2011.00634.x

Valeski, T. N., and Stipek, D. J. (2001). Young children's feelings about school. *Child Dev.* 72, 1198–1213. doi: 10.1111/1467-8624.00342

Valiente, C., Lemery-Chalfant, K., Swanson, J., and Reiser, M. (2008). Prediction of children's academic competence from their effortful control, relationships, and classroom participation. *J. Educ. Psychol.* 100:67. doi: 10.1037/0022-0663.100.1.67

Walker, H. M. (2004). Commentary: Use of evidence-based interventions in schools: Where we've been, where we are, and where we need to go. *Sch. Psychol. Rev.* 33, 398–407. doi: 10.1080/02796015.2004.12086256

Walker, S. (2005). Gender differences in the relationship between young children's peer-related social competence and individual differences in theory of mind. *J. Genet. Psychol.* 166, 297–312. doi: 10.3200/GNTP.166.3.297-312

Walker, S., Berthelsen, D., and Irving, K. (2000). Teacher reports of social behavior and peer acceptance in early childhood: Sex and social status differences. *Paper presented at the Australian Research in Early Childhood Education Conference, Canberra.*

Walker, S., Irving, K., and Berthelsen, D. (2002). Gender influences on preschool children's social problem-solving strategies. *J. Genet. Psychol.* 163, 197–209. doi: 10.1080/00221320209598677

Watson, A. C., Nixon, C. L., Wilson, A., and Capage, L. (1999). Social interaction skills and theory of mind in young children. *Dev. Psychol.* 35:386. doi: 10.1037/0012-1649.35.2.386

Wellman, H. M. (2014). *Making Minds: How Theory of Mind Develops*. Oxford University Press.

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

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

Wentzel, K. R. (1999). Social-motivational processes and interpersonal relationships: Implications for understanding motivation at school. *J. Educ. Psychol.* 91:76. doi: 10.1037/0022-0663.91.1.76

Wentzel, K. R. (2005). "Peer relationships, motivation, and academic performance at school," in *Handbook of competence and motivation*, eds A. J. Elliot, C. S. Dweck, and D. S. Yeager (New York, NY: The Guilford Press), 279–296.

Westra, E., and Carruthers, P. (2018). Theory of mind. *Encyclop. Evol. Psychol. Sci.* 69, 71–76.

Wood, J. J. (2007). Academic competence in preschool: Exploring the role of close relationships and anxiety. *Early Educ. Dev.* 18, 223–242. doi: 10.1080/10409280701282868

Wu, Y., Wu, J., Chen, Y., Han, L., Han, P., Wang, P., et al. (2015). Shyness and school adjustment among Chinese preschool children: Examining the moderating effect of gender and teacher–child relationship. *Early Educ. Dev.* 26, 149–166. doi: 10.1080/10409289.2015.970503

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 30 September 2024
DOI 10.3389/fpsyg.2024.1472538

OPEN ACCESS

EDITED BY
Ann Dowker,
University of Oxford, United Kingdom

REVIEWED BY
Schnell Zsuzsanna,
University of Pécs, Hungary
Cesco Willemse,
University of Twente, Netherlands

*CORRESPONDENCE
Jelena Ristic
jelena.ristic@mcgill.ca

RECEIVED 29 July 2024
ACCEPTED 17 September 2024
PUBLISHED 30 September 2024

CITATION
Mayrand F, Capozzi F and Ristic J (2024) Gaze communicates both cue direction and agent mental states.
Front. Psychol. 15:1472538.
doi: 10.3389/fpsyg.2024.1472538

Gaze communicates both cue direction and agent mental states

Florence Mayrand[1], Francesca Capozzi[2] and Jelena Ristic[1]*

[1]Department of Psychology, McGill University, Montreal, QC, Canada, [2]Department of Psychology, Université du Québec à Montréal, Montreal, QC, Canada

Introduction: Although it is well established that humans spontaneously attend to where others are looking, it remains debated whether this gaze following behavior occurs because gaze communicates directional information (i.e., *where* an agent is looking) or because gaze communicates an agent's inferred mental content (i.e., *what* an agent perceives), both of which rely on the processes involved in the general Theory of Mind ability.

Methods: To address this question, in two Experiments we used a novel task to measure how spatially dissociated and spatially combined effects of an agent's gaze direction and perceived mental content influence target performance. We also contrasted performance for social directional cues and nonsocial arrows.

Results: Our data revealed that performance was compromised when cue direction and mental content dissociated relative to when they combined. Performance for dissociated components was especially prominent when a social avatar served as a cue relative to a comparison arrow.

Discussion: Together, these data show that a typical gaze signal communicates information about both *where* an agent is attending and *what* they are attending to.

KEYWORDS

gaze direction, perspective taking, cue direction, mental content, theory of mind

1 Introduction

Visual information conveyed by gaze (eyes, head, or body deviation) enables quick communication of social messages (Capozzi and Ristic, 2018). As such, the ability to follow gaze has been implicated in both basic social functions like gaze following and joint attention (Frischen et al., 2007) as well as more complex social behaviors such as attitude formation (Toscano et al., 2018) and social status inference (Capozzi et al., 2019). Strikingly however, the large body of research on gaze following remains agnostic on the nature of messages conveyed by gaze (Capozzi and Ristic, 2020; Frischen et al., 2007). Specifically, there remains a key outstanding question of whether humans follow gaze because it conveys directional information about *where* the items of a gazer's interest are located or because gaze also coveys mentalistic information about *what* the gazer perceives (Capozzi and Ristic, 2020). Using a novel behavioral test procedure, here we dissociate the contributions of directional and mentalistic components of the gaze signal and in doing so, show that decoupling of this information leads to the largest detriments in performance. This suggests that ordinarily gaze communicates *both* directional information about where the gazer is looking and mentalistic information about *what* the gazer is perceiving.

The nature of signals communicated by gaze has been the subject of a longstanding debate (e.g., Capozzi and Ristic, 2020). The proponents of the directionality account argue that spontaneous following of gaze, often experimentally demonstrated by the gaze cuing procedures (Friesen and Kingstone, 1998) or dot perspective task (Samson et al., 2010), reflects the influence of the cue's direction indicating where in space the gazer's attention is directed

(Cole and Millett, 2019; Santiesteban et al., 2017). In contrast, the proponents of the mentalizing account maintain that this behavior is instead driven by the spontaneous adoption of the gazer's visual perspective, which aids with observers inferring and sharing the representation of the gazed-at object with the gazer (Apperly and Butterfill, 2009). Supporting the directionality account, gaze following has been found to occur similarly regardless of whether it is elicited by the direction of social gaze or a nonsocial cue (e.g., an arrow; Kingstone et al., 2019). Supporting the mentalizing account, gaze following magnitude is reduced when participants believe gaze cues are delivered by a randomized computer sequence (Wiese et al., 2012) or a person whose line of sight is obstructed (Baker et al., 2016).

Detecting and understanding gaze signals is often understood as a key precursor and a component of Theory of Mind, which generally denotes one's ability to take the mental perspective of others (Premack and Woodruff, 1978; Baron-Cohen, 1991). Within this context, directional and mentalistic aspects of gaze track with two proposed levels of mental perspective taking complexity, thus offering insights into how each of the gaze processing aspects may inform the Theory of Mind (ToM). Specifically, the directional gaze account appears to align with the processes associated with Level-1 visual perspective taking. This aspect of gaze communication requires the observer to understand the physical orientation of the gaze in others, which is foundational in both social cognition and ToM (Kessler and Rutherford, 2010). Crucially, Level-1 visual perspective taking involves the ability to comprehend whether an object is visible within the line of sight of another person, regardless of whether it is visible from one's own line of sight (Flavell et al., 1981). Comparatively, the mentalizing gaze account appears to align with the processes associated with Level-2 visual perspective taking, in which one understands the visual aspects of a scene relative to an imagined viewpoint of a gazer, such that one is able to understand that even if an object is visible to the self and another person, this does not mean they share the same mental representation (Flavell et al., 1981). This requires representation of the gazer's mental states thereby invoking a deeper level of mental state attribution (Todd et al., 2017). To maintain connection with the existing research and the ongoing dialogue between directional and mentalistic accounts (e.g., Capozzi and Ristic, 2020), here we refer to the two processes implicated in gaze signals as the directional component of gaze and the mentalistic component of gaze, whereby the directional component is conceptualized as reflecting the influence of gaze cue direction on behavior (i.e., Level-1 perspective taking) while the mentalizing component is conceptualized as reflecting the influence of the understanding of the gazer's mental perspective on behavior (i.e., Level-2 perspective taking).

To investigate the contribution of each of these components in gaze following, we designed a novel behavioral task in which we measured the combined and dissociated effects of these two processes on target performance. We ran two experiments in which participants were presented with an image of a central cue (a human avatar or an arrow) and were asked to localize a peripheral target that was flanked by a non-target distractor (Figure 1). Providing directional information to the observer, the cue indicated a left or right spatial location, that is, it either pointed toward the response target or toward a non-target distractor equally often. Providing mentalistic information, the avatar also "perceived" either same target as the observer or the non-target distractor due to our manipulation of targets which afforded multiple representations depending on the point of view. Critically, in the combined conditions, the directional and mentalistic processes were congruent, as they would be in a typical gaze signal, such that the cue both indicated the target directionally the observer gazed to and mentally perceived its content from its perspective. Critically, in the dissociated condition, they could directionally look at the same target identity as the observer, but they may nevertheless perceive a different object. Similarly, the cue may also look at a different target identity as an observer.

Two Experiments were run. They were identical except that in Experiment 1 social gaze and nonsocial avatar were manipulated between subjects while in Experiment 2 the two cues were manipulated within participants. Experiment 2 thus also provided a direct replication of the initial proof of concept obtained in Experiment 1.

If gaze typically communicates information about both cue direction and mental content of the gazer, target performance should suffer the most when these two components dissociate relative to when they combine. That is, a condition in which the avatar is looking at a peripheral location containing the observer's target but perceives a distractor from its own perspective should result in slower responses relative to a condition in which the avatar is looking at a peripheral location containing the observer's target and perceives the same target from its own perspective. This is because in the former case, the directional and mentalistic information communicated by the avatar are spatially dissociated while in the latter case the two pieces of information are spatially combined as they would be in a typical gaze signal. Hence, when directional and mentalistic processes occur together in a congruent manner (i.e., the avatar is both looking at and perceiving the same target as the observer from its visual perspective) performance is facilitated. To understand whether any these effects may be unique to directional cues delivered by social agents, we also subjected a nonsocial central arrow cues to the same test.

2 Experiment 1

While gaze is typically understood to reflect human social communication, directional cues such as arrows have often been used as a comparison (i.e., Santiesteban et al., 2014; Kingstone et al., 2019; Kuhn and Kingstone, 2009), due to their similar directional representation but lesser social value. In Experiment 1, we examined combined and dissociated contributions of the mentalistic and directional components of gaze elicited by a social avatar and nonsocial arrow cues. If gaze typically invokes both directional and mentalistic representations in observers, target-related response detriments for dissociated representations should be more pronounced when a social avatar serves as a cue relative to when a nonsocial arrow serves as cue.

2.1 Methods

2.1.1 Participants

An *a priori* power analysis using an estimated moderate size estimate of $r = 0.25$ for the variability in the magnitude of the gaze cuing with mental state attribution (approximated from Moriguchi et al., 2006; Sulpizio et al., 2015; Tomei et al., 2017) indicated that data from about 95 participants would yield power of 0.8 and data from 130 participants would yield power of 0.9 (Alpha = 0.05). Data from

220 participants were included in the analysis[1], with 111 participants randomly assigned to view the avatar cue (97 women,14 men, Mean age = 20.53 years, *SD* = 2.34) and 109 participants (92 women, 15 men, 2 other, Mean age = 20.67 years, *SD* = 2.82) randomly assigned to view the arrow cue. All procedures were approved by the University's Research Ethics Board. Informed consent was obtained from all participants. Participants were recruited from the McGill University participant pool and compensated with course credit[2]. All participants reported native English fluency, no history of psychiatric or neurological conditions, and normal or corrected-to-normal vision. Deidentified data are available at osf.io/3xcqk.

2.1.2 Apparatus and stimuli

Figure 1 illustrates the stimuli (A), sample trial sequence (B), and the four test conditions for each avatar and arrow cue (C). Cues were images of a human avatar and an arrow, which were equated for length (Figure 1A). The cues were positioned at fixation. Target and distractor stimuli, which were shown on the left and right of fixation respectively, were a letter E, a number 3, and a number 8 (Figure 1C). They were equated for size. Each response target, either 3, E, or 8, was always flanked by a unique distractor, creating three unique target-distractor combination (3 / E; 3 / 8; E / 8).

The study was administered online via Testable (https://www.testable.org/). The experiment was launched on participants' personal computers. Because participants were alone in their own environments to complete the study, they were instructed to minimize distractions in their surroundings for the duration of the experiment. The entire stimulus response display image including all stimuli in proportions was scaled to approximately 50% of individual participants' screens. The display target images were rendered in grayscale; the central cues were rendered in green.

2.1.3 Design

The experiment was a mixed design, with *Cue Type* (Avatar; Arrow) included as a between subjects variable, and *Cue Direction Consistency* (Consistent; Inconsistent) and *Target Content Consistency* (Consistent; Inconsistent) included as within subjects variables.

Cue Type manipulated the type of the central cue and varied between an avatar and an arrow. Half of the participants were presented with an avatar cue and the other half was presented with an arrow cue. The cue assignment was randomized across participants.

Cue Direction Consistency manipulated whether the response target appeared at the spatial location directionally indicated by the cue (i.e., Consistent; Figure 1C) or at the opposite location (i.e., Inconsistent). This variable is a composite of cue direction and target location factors, each of which varied independently and equiprobably between left and right spatial positions.

Target Content Consistency manipulated the observed and inferred avatar's mental content. The avatar's mental content (i.e., what they would perceive from their viewpoint) was either consistent or inconsistent with the participant's perception of target identity. Consistent target content between the observer and the avatar was induced using target stimuli that invoke the same percept from the observer's and the avatar's perspectives (i.e., number 8; Figure 1C). That is, the response target 8 would have the same percept from both the observer's and the avatar's perspective. Inconsistent target content between the observer and the avatar was induced using target stimuli that are perceived as a target from the observer's perspective but a distractor from avatar's perspective and vice versa. That is, for the avatar, the observer's response target 3 would appear as a distractor letter E while the response target E would appear as a distractor number 3. Target identity and target location were intermixed, varied independently and equiprobably with cue direction. Each target-distractor combination appeared equally often.

The combination of these variables created four key conditions of interest. In the *Consistent direction/Consistent content (CC)* and the *Inconsistent direction/Inconsistent content (II)* conditions, cue direction and target mental content combined, such that both were either consistent (CC) or inconsistent (II). That is, the cue either pointed at or away from the target, which matched the observer's and the avatar's perspective. In contrast, in the *Consistent direction/Inconsistent content (CI)* and the *Inconsistent direction/Consistent content (IC)* conditions, cue direction and target mental content dissociated. In the CI condition, the avatar looked the response target but perceived a distractor (i.e., the observer's response target 3 appears as a distractor E). In contrast, in the IC condition, the avatar looked at a distractor but perceived the target at that location, since the non-target distractor may be perceived as the target from the avatar's perspective (i.e., distractor 3 appears as the response target E).

2.1.4 Procedure

Example trial sequence is shown in Figure 1B. Trials began with an image of a response target (3, E, 8) for 500 ms to inform participants about the target they need to respond to for that trial. Then, the display showing the central cue and the target/distractor combination was presented. The display remained visible for 2000 ms or until participants responded. Participants were informed that the cue direction was not predictive of target location or target identity. They were asked to localize the target quickly and accurately by pressing the 'b' and 'h' keys on the keyboard. Target location (left, right) and key response ('b' or 'h') pairing was counterbalanced between participants.

The experiment consisted of four blocks of 80 trials, for a total of 320 trials. Each block contained 20 trials for each of the 4 test conditions (CC, II, CI, IC). All conditions were intermixed and presented equally often using a random sequence. Eight practice trials with performance feedback were run first. The study was completed in 38.85 min on average (SD = 9.13 min).

1 A total of 275 participants completed the study. Participants achieving less than 80% overall accuracy (i.e., over 20% of trials lost due to anticipations, timeouts, or incorrect responses, e.g., McCrackin and Itier, 2019; Van Selst and Jolicoeur, 1994) were excluded from analyses (n=55).

2 Participants also completed the 28-item Interpersonal Reactivity Index (IRI; Davis, 1983), where, for each question (e.g., "I try to look at everybody's side of a disagreement before I make a decision"), they responded using a 5-point Likert scale ranging from 1 - "Does not describe me well" to 5 - "Describes me very well." Exploratory correlation analyses linking individual RT performance in each of the four test conditions (CC, II, CI, IC) suggested that there was a positive relationship between the IRI score and RTs in the II (r=0.189; p=0.047) and CC (r=0.186; p=0.051) conditions in the Avatar group only (Avatar group: 0.111<rs<0.189, 0.246<p<0.047; Arrow group: −0.023<rs<−0.059, 0.545<p<0.808).

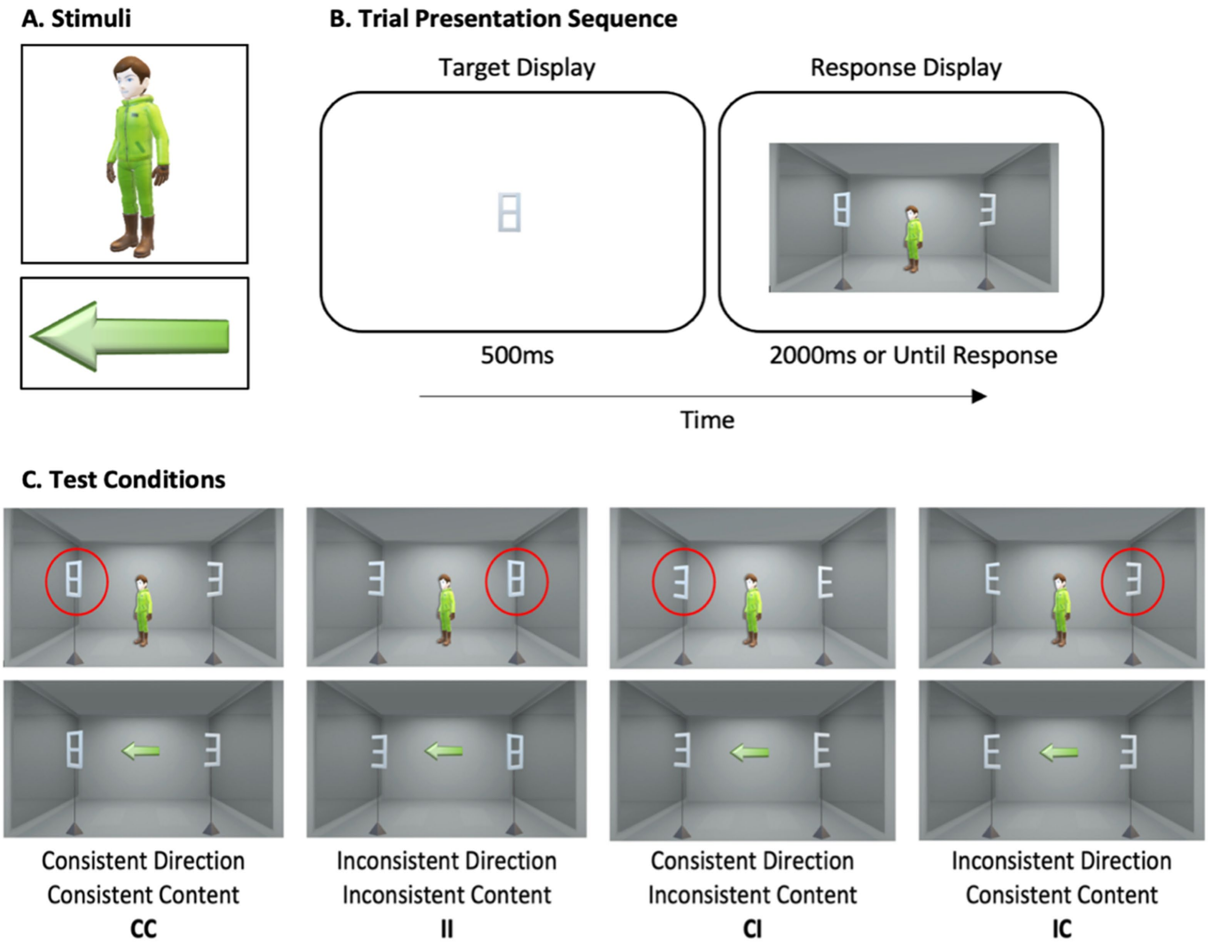


FIGURE 1
Stimuli, example stimulus presentation sequence, and test conditions. **(A)** Illustration of cue and target stimuli. **(B)** Example trial sequence. Participants were first presented with an image of the trial response target for 500 ms. The response display was then presented for 2000 ms or until participants responded. **(C)** Illustration of the four test conditions for each cue type. Target is highlighted for the reader with a red circle. Stimuli are not drawn to scale.

2.2 Results

Response anticipations (RT < 200 ms) and timeouts (RT > 1800s) accounted for 1.49% of all trials and were removed from analyses. Performance was high overall at 92.22%. Accuracy was examined across *Cue Type* (Avatar, Arrow; between-subjects) and the four test *Conditions* (CC, II, CI, IC; within-subjects) using a mixed-effects ANOVA. No speed-accuracy tradeoff was evident in the data, with CC ($M = 0.96$) and II ($M = 0.96$) conditions returning higher overall accuracy then CI ($M = 0.89$) and IC ($M = 0.88$) conditions, [F (2.30, 502.11) = 303.67, $p < 0.001$, MSE = 0.002, $\eta p^2 = 0.58$]. *Cue Type x Condition* interaction was not reliable, [$F(3, 654) = 0.19$, $p = 0.90$, $MSE = 0.002$, $\eta p2 = 0.001$]. When Mauchly's test was significant, the Greenhouse–Geisser degrees of freedom are reported. All follow-up, t-tests two-tailed and Bonferroni corrected. Data were analyzed using SPSS 27.

Mean correct Response time was analyzed using a mixed-effects ANOVA with *Cue Type* (Avatar; Arrow) included as a between-subjects variable, and *Cue Direction Consistency* (Consistent; Inconsistent) and *Target Content Consistency* (Consistent; Inconsistent) included as within-subjects variables. Figure 2 illustrates mean RTs as a function of Cue type, Cue Direction consistency and Target Content consistency.

The analysis indicated that all main effects were significant. A main effect of *Cue Type* (1, 218) =7,789, $p < 0.0001$, MSE = 53,165, $\eta p^2 = 0.973$, indicated overall faster responses to targets cued by the arrow than targets cued by gaze. A main effect *Cue Direction Consistency* indicated that overall CC trials were faster than II trials, $F(1, 218) = 11.87$, $p < 0.001$, $MSE = 788.22$, $\eta p^2 = 0.052$, which replicates the typical cue directionality gaze following effects [e.g., Friesen and Kingstone (1998); see Capozzi and Ristic (2018) for recent review of this literature]. Finally, a main effect of *Target Content Consistency* indicated that overall consistent mental content between the observer and the cue generated faster response times than inconsistent mental content, $F(1, 218) = 13.10$, $p < 0.001$, $MSE = 660.19$, $\eta p^2 = 0.057$.

A two way interaction between *Cue Direction Consistency* x *Target Content Consistency* was significant, $F(1, 218) = 1214.10$, $p < 0.001$, $MSE = 4748.76$, $\eta p^2 = 0.85$, and indicated that consistent direction trials were responded to faster when paired with consistent mental content relative to when paired with inconsistent target content [CC vs. CI for arrow: $t(108) = -24.26$, $p < 0.001$; for avatar: $t(110) = -23.73$, $p < 0.001$]. Similarly, inconsistent direction trials were responded to faster when

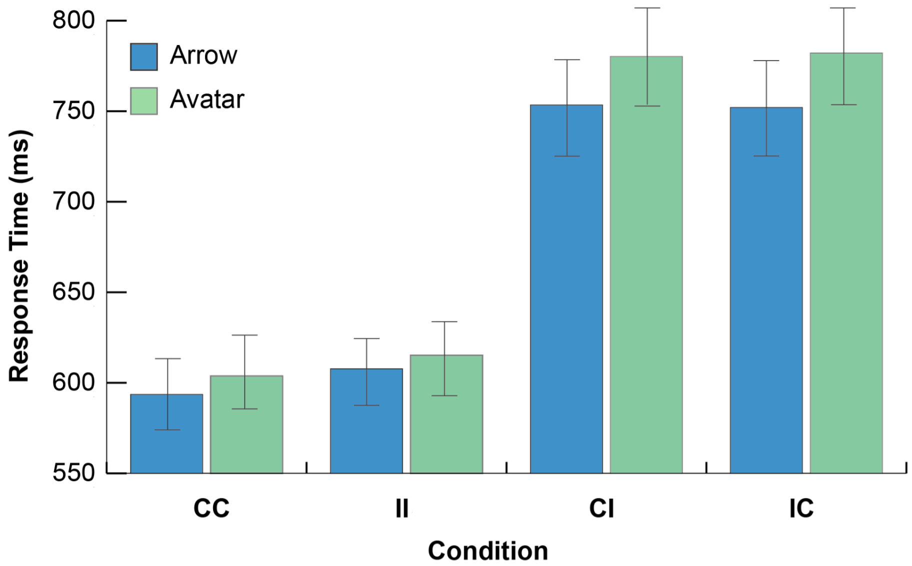


FIGURE 2
Experiment 1 results. Mean response time as a function of cue type, cue direction consistency, and target content consistency conditions (CC, consistent cue direction/consistent target content; II, inconsistent cue direction/inconsistent target content; CI, consistent cue direction/inconsistent target content; IC, inconsistent cue direction/ consistent target content). Error bars depict the 95% CI.

paired with inconsistent mental content relative to when paired with consistent mental content [IC vs. II for arrow: $t(108) = -21.80$, $p < 0.001$; for avatar: $t(110) = -22.77$, $p < 0.001$].

Finally, and supporting our hypothesis, a significant three-way interaction between *Cue Type*, *Cue Direction Consistency*, and *Target Content Consistency*, $F(1, 218) = 4.48$, $p = 0.035$, $MSE = 4748.76$, $\eta p^2 = 0.02$, indicated slower performance for the avatar cue relative to the arrow cue for the CI (780 ms avatar vs. 753arrow) and IC ($M = 782$ ms avatar vs. 752 ms arrow) conditions, while there was less of a cue type difference for the CC (604 ms Avatar vs. 594 ms Arrow) and II (615 ms Avatar vs. 608 ms Arrow) conditions. Follow-up independent group t-tests comparing the mean RTs for the Avatar and Arrow cue CI [$t(218) = 1.48$, $p = 0.14$] and IC [$t(218) = 1.64$, $p = 0.10$] conditions did not reach significance. No other effects or interactions were reliable (*Cue Type x Cue Direction Consistency*; *Cue Type x Target Content Consistency*, both $Fs < 1$).

2.3 Discussion

Experiment 1 data provided proof of concept evidence indicating a performance detriment for targets following dissociated directional and mentalistic aspects of the gaze cue (IC and CI conditions) relative the combined ones (CC and II conditions). Further, and in line with the idea of gaze cues transmitting primarily social information, this behavioral performance detriment was larger when a social avatar served as a cue relative to when a nonsocial arrow served as a cue. Our data also confirmed that participants responded overall faster to targets that are directionally cued (vs. those that are not, demonstrating a classic cueing effect; Friesen and Kingstone, 1998; Frischen et al., 2007) and to those targets with matched mental content. Thus, Experiment 1 results provide one of the first pieces of experimental evidence showing that gaze signals convey information both about *where* the gazer is looking and *what* they are looking at.

However, due to the nature of our design in which they key difference between a social and nonsocial attribution were made based on a between group differences, it remains unknown if those differences reflect differences in social cue processing versus differences between the two groups of participants. To address this issue, and to provide a full replication of our initial finding, in Experiment 2, we ran a preregistered high-powered replication of Experiment 1.

3 Experiment 2

Experiment 2 was identical to Experiment 1 except that we recruited new group of participants and manipulated cue type within participants. As before, we hypothesized that if gaze signal includes both directional and mentalistic information, target-related response detriments for dissociated directionality and mental content should be more pronounced when social avatar serves as a cue relative to when a nonsocial arrow serves as a cue.

3.1 Methods

The study was pre-registered (osf.io/fwptb). Deidentified data are available at osf.io/3xcqk.

3.1.1 Participants, apparatus, stimuli, design, and procedure

An *a priori* power analysis using an estimated moderate size estimate of $r = 0.25$ for the variability in the magnitude of the gaze cuing with mental state attribution (approximated from Moriguchi et al., 2006; Sulpizio et al., 2015; Tomei et al., 2017) indicated that data from about 95 participants would yield power of 0.8 and data from 130 participants would yield power of 0.9 (Alpha = 0.05).

10.3389/fpsyg.2024.1472538

The data from 136 new participants (114 women, 22 men; mean age = 20.51 years, SD = 1.34) were analyzed[3]. Participants were recruited from the McGill University participant pool and compensated with course credit. All procedures were approved by the University's Research Ethics Board. Informed consent was obtained from all participants. All participants reported native English fluency, no history of psychiatric or neurological conditions, and normal or corrected-to-normal vision[4].

Stimuli, apparatus, design, and procedure were identical as in Experiment 1 except that (i) all participants responded to targets cued by the avatar and the arrow cue, (ii) *Cue Type* was blocked, such that half of the blocks presented an avatar cue, and the other half presented the arrow cue. First block assignment was randomized, and subsequent order alternated; (iii), due to an addition of a variable, the total number of trials increased with participants completing eight blocks of 64 trials, for a total of 512 trials. Each block contained 16 trials for each of the four test conditions (CC, II, CI, IC). All conditions were intermixed and presented equally often using a random sequence. Sixteen practice trials with performance feedback were run first. The study was completed in 62.14 min on average (SD = 14.65 min).

3.2 Results

Response anticipations (RT < 200 ms) and timeouts (RT > 1800s) accounted for 1.99% of trials and were removed from analyses. When Mauchly's test was significant, the Greenhouse–Geisser degrees of freedom are reported. All follow-up paired, two-tailed t-tests were Bonferroni corrected. Data were analyzed using SPSS 27.

Overall, the task was well done with 93.03% response accuracy. Mean accuracy was examined for each *Cue Type* (Avatar, Arrow) and *Condition* (CC, II, CI, IC) using a repeated measures ANOVA. The analyses indicated no speed-accuracy trade-off, as the CC ($M = 0.97$) and II ($M = 0.96$) conditions were overall responded to more accurately (and faster, as described in Results) than the CI ($M = 0.90$) and IC ($M = 0.90$) conditions [*Condition* main effect, $F(2.39, 323.23) = 208.51$, $p < 0.0001$, MSE = 0.002, $\eta p^2 = 0.61$] for both cue types. Main effect of *Cue Type* ($F < 1$) and the *Cue Type x Condition* interaction, $F(2.75, 370.99) = 2.63$, $p = 0.055$, MSE = 0.001, $\eta p^2 = 0.019$, were not reliable.

To remind, we reasoned that if gaze communicates both directional and mentalistic content, a detriment in target performance should emerge when these two components of a gaze signal are dissociated in the CI and IC conditions. We also predicted that this performance detriment would be more pronounced when the social avatar cue.

A repeated measures ANOVA with *Cue Type* (Avatar; Arrow), *Cue Direction Consistency* (Consistent; Inconsistent), and *Target Content Consistency* (Consistent; Inconsistent) was used to examine mean correct RTs. These means are plotted in Figure 3 as a function of Cue type, Cue Direction consistency, and Target Content consistency conditions.

The results supported our predictions by returning a reliable two-way interaction between *Cue Direction* and *Target Content Consistency* [$F(1, 135) = 681.68$, $p < 0.001$, $MSE = 10580.12$, $\eta p2 = 0.835$] as well as a reliable three-way interaction between *Cue Type*, *Cue Direction Consistency* and *Target Content Consistency* [$F(1, 135) = 12.914$ $p < 0.001$, MSE = 12111.288, $\eta p^2 = 0.087$]. The first interaction indicated that, overall, targets cued by consistent cue direction were responded to faster when that cue direction was paired with consistent target content relative to when it was paired with inconsistent target content [CC vs. CI; Avatar: $t(135) = -24.92$, $p < 0.001$; Arrow: $t(135) = -22.670$, $p < 0.001$]. Likewise, uncued targets (i.e., inconsistent direction trials) were responded to faster when paired with inconsistent target content relative to when paired with consistent target content [II vs. IC; Avatar: $t(135) = -23.84$, $p < 0.001$; Arrow: $t(135) = -22.25$, $p < 0.001$]. This replicates data from Experiment 1.

Further, a significant three-way interaction between *Cue Type*, *Cue Direction Consistency*, and *Target Content Consistency* [$F(1, 135) = 12.914$ $p < 0.001$, MSE = 12111.288, $\eta p^2 = 0.087$]. indicated that there was a larger performance difference between the dissociated CI and IC conditions, in which the cue directionality and target mental content are dissociated [CI: $t(135) = -3.126$, $p = 0.002$; IC: $t(135) = -3.088$, $p = 0.002$], when the social avatar served as a cue relative to when a nonsocial arrow served as a cue. In other words, responses to targets following gaze cues were slower than responses to targets following arrow cues in conditions in which directionality and mental content are dissociated. Thus, the dissociation of the two components induced a larger performance determent during avatar trials than during arrow trials. In contrast, there was no difference in response to the targets across the two cue types in the CC and II conditions, where cue directionality and target mental content combined [CC: $t(135) = -1.756$, $p = 0.081$; II: $t(135) = -0.298$, $p = 0.766$; *Cue Type x Cue Direction Consistency*, $F(1,135) = 1.501$, $p = 0.223$]. This replicates our results from Experiment 1, and a well-known finding in the literature showing that social and nonsocial cues typically induce similar directionality effects (Ristic et al., 2002). The interaction between *Cue Type* and *Target Content Consistency* was not reliable ($F < 1$).

As before, the ANOVA returned reliable main effects of *Cue Type* [$F(1, 135) = 7.59$, $p = 0.007$, $MSE = 4531.75$, $\eta p^2 = 0.053$] and *Cue Direction Consistency* [$F(1, 135) = 7.951$, $p = 0.006$, $MSE = 5366.289$, $\eta p^2 = 0.056$], with overall faster responses to targets cued by the arrow's direction than targets cued by the avatar's gaze, and overall faster responses in the CC condition than the II condition across both cue types (i.e., an overall gaze following effect). A main effect of *Target Content Consistency* approached significance, $F(1, 135) = 3.815$, $p = 0.053$, $MSE = 858.22$, $\eta p^2 = 0.027$, with trials on which the target content was consistent with the observer's mental representation generating numerically lower RTs than trials on which the target content representation diverged across the observer and the cue.

3 A total of 162 undergraduate students completed the study. Data from 26 participants were removed based on the pre-registered exclusion criteria of having less than 80% overall accuracy (i.e., over 20% of trials for each participant lost due to response anticipations, timeouts, or incorrect responses, e.g., McCrackin and Itier, 2019; Van Selst and Jolicoeur, 1994).

4 As in Experiment 1, participants completed the 28-item Interpersonal Reactivity Index (IRI; Davis, 1983). Exploratory correlation analyses linking individual RT performance in each of the four test conditions (CC, II, CI, IC) with the overall IRI score for yielded no reliable effects ($-0.004 <$ *rs* < -0.088, all *ps* > 0.308).

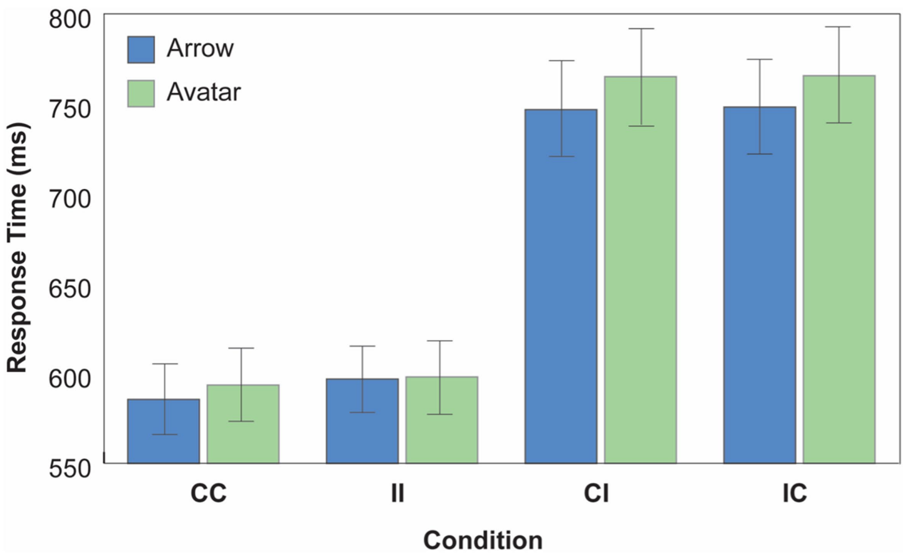


FIGURE 3
Experiment 2 results. Mean correct response time (RT) as a function of cue type, cue direction consistency and target content consistency conditions (CC, consistent direction/consistent content; II, inconsistent direction/inconsistent content; CI, consistent direction/inconsistent content; IC, inconsistent direction/consistent content). Error bars depict the 95% CI.

Thus, to summarize, in line with our predictions, in Experiment 2, we once again found a reliable detriment in target performance when cue directionality and target mental content (IC and CI conditions) are experimentally dissociated relative to when they are combined (CC and II conditions). Further, and critically, this behavioral performance detriment was larger when a social avatar served as a cue relative to when a nonsocial arrow served as a cue. As such, this result replicated Experiment 1 to once again support the notion that social gaze normally conveys information both about *where* the gazer is looking and *what* they are looking at, i.e., invoking both Level 1 and Level 2 visual perspective processes. When these processes are combined, typical effects of facilitated performance for gaze-directed targets are found. When they are dissociated, a large slowdown in performance is observed.

4 General discussion

Grasping the nature of messages conveyed by human gaze is important for understanding the means and the underlying mechanisms of our nonverbal social communication. Using a novel paradigm that measures both combined and dissociated contributions of gazer's cue directionality and their mental content, in two experiments, here we showed that when these two aspects of the gaze signal are dissociated, significant slowing in performance is found. This is in contrast to the conditions in which cue direction and avatar's mental content were congruent, where performance was reliably facilitated by such combined representation. Furthermore, the detriment in performance was more pronounced when social avatar delivered the cues. As such, these results are one of the first experimental demonstrations of the influence of both cue directionality *and* mental content in gaze communication and provide a window into the cognitive processes involved in visual perspective taking associated with the Theory of Mind. That is, our findings suggest that the gaze signal typically communicates information about both cue direction and the gazer's visual perspective, i.e., both *where* agents are looking and *what* they perceive. In other words, understanding of the line of sight depends both on the computation of the direction of the cue and the inferred mental perspective of the gazer. We next bring up and discuss two potential implications of this finding.

First, the typical gaze following response appears to reflect the computation of both the gaze cue's direction, i.e., the line of sight *and* the representation of the gazer's mental content. This conclusion follows from our finding showing impaired performance in conditions in which the cue's directionality and target mental content were dissociated. That is, target performance was slowed in a similar manner by both the absence of the cue's directional content *and* the absence of the mental representation of the target. Thus, in typical gaze following responses, both cue's directional information and agents' mental state appear to be relayed in a joint fashion. When these two processes are experimentally dissociated (i.e., in CI and IC conditions), the gaze following response is slowed. This finding provides an empirical resolution to the longstanding debate in the field of whether gaze conveys directional information or mental content, showing the involvement of both the domain general processes linked to processing of cue directionality as well as domain specific processes linked to computations of social mental content of the gazer (Capozzi and Ristic, 2020).

Dovetailing with this point, our results also showed that dissociating the gaze cue direction information from its mental content signal impacted target performance the most when the social avatar relative to a nonsocial arrow served as a cue. Thus, it seems likely that while the components of cue direction and mental content may be combined in social signals, they appear less intertwined and thus more easily dissociated in nonsocial signals (Marotta et al., 2012). For example, while a simple "line of sight" computation may be easily attributed to a directional cue such as an arrow, this nonsocial stimulus does not possess a mental or visual perspective. That said, while it may be counterintuitive to attribute 'mental state' representation to a nonsocial cue, a joint representation of cue directionality and its meaning may still occur in this condition but to a lesser extent given that behaviorally relevant cues like arrows may convey highly meaningful messages in daily life (e.g., Santiesteban et al., 2014). Indeed, recent studies argue that the differences between "social" and "nonsocial" stimuli may reflect both implementation as well as mechanisms differences and could vary with tasks and environmental contextual conditions (Lockwood et al., 2020). The difference between the magnitudes of the dissociated effects between social and nonsocial cues thus may reflect the increased difficulty in disjointing the combined representation of directionality and meaning for social vs. overlearned nonsocial cues (e.g., Ristic and Kingstone, 2012). Future studies are needed to examine how dissociating cue direction from mental content representation may be affected by the differences in the cues' social and learned values.

More generally, these findings provide significant insights into the cognitive processes associated with Theory of Mind by demonstrating that gaze signals include information on both *where* an agent is looking (directional information) and *what* they are perceiving (mental content). This joint communication maps well onto the two levels of visual perspective-taking. Specifically, the directional component of gaze corresponds to Level 1 perspective-taking, where the observer understands the physical orientation of the gazer's attention and comprehends that something is within the gazer's line of sight. In contrast, the mentalistic component aligns with Level 2 perspective-taking, which involves understanding the visual scene from the gazer's viewpoint, thus requiring a more complex mental state attribution. This implies that Theory of Mind involves not only the ability to infer others' mental states but also the ability to integrate those inferences with directional information within the environment. Furthermore, our results support the idea that Theory of Mind operates under a hierarchy of cognitive abilities, such that low-level processing of directional cues and higher-level processing of mental states both relate to Theory of Mind processes (Jarrold et al., 2000; Happé and Frith, 2006). Understanding these mechanisms enhances our knowledge of the cognitive processes underlying Theory of Mind, emphasizing a native link between basic directional computation and more involved mental state attribution in social communication (Qureshi and Monk, 2018). Our results suggest that disruptions in these processes may be associated with social cognitive deficits observed in various neurodevelopmental conditions such as autism spectrum disorders, where Theory of Mind tends to be difficult (Rosello et al., 2020; Baron-Cohen et al., 1985; Kimhi, 2014).

Although our work provides valuable new insights about social signals, there are several outstanding points worth considering in future extensions of this work. First, one might wonder if the reported detriment in performance for the dissociated conditions may reflect visual differences in targets, since response targets in the dissociated conditions afford multiple mental representations (i.e., 3 and E) while response targets in the combined condition (i.e., 8) afford a single mental representation. It is important to note that our results showing a larger detriment in performance in dissociated conditions for an avatar relative to an arrow cue occurred across same targets. The same performance detriment also occurred across two experiments and was stable in comparison between and within participants. Thus, while the ease of target representation may overall facilitate performance as indicated in our data, the differences in representation across the two cue types, which are critical for our conclusions, persisted across targets that afford single and multiple mental representations. Nevertheless, it would be important to understand in future work if variability in mental representation of targets influences the contributions of directionality vs. mental content to gaze signals. Second, and dovetailing with this first point, additional future challenge concerns assessments of perspective-taking, given that differentiating between the influences of egocentric and altercentric intrusions, as indices of perspective taking, in tests like these is complex (i.e., Del Sette et al., 2022). Although we aimed to study the impact of cue directionality and mentalistic information on observers' target judgments, it is possible that egocentric and altercentric intrusions still occurred in our data, where participants made judgments from the avatar's directional or mental perspective. As a case in point, our data indicated an altercentric intrusion which was evident in the CI condition where observers are spontaneously influenced, i.e., slowed down in responses, by the mismatch between their own and avatar's visual perspective of target identity. Future studies are needed to understand how assuming self vs. other perspective may influence these attributions and whether the ease of adopting different visual perspectives (e.g., spontaneous vs. explicit) may affect the interpretation of directional and mentalistic aspects of gaze. Finally, it would be important to theoretically link Level-1 and Level − 2 perspective with the respective processing of cue information.

In sum, using a novel behavioral paradigm, across two Experiments, here we show experimentally that that gaze signal conveys both information about *where* an agent is looking and *what* they perceive. This finding provides a new perspective on social signaling by highlighting a complex interplay between the processes involved in contributions of gaze directionality and mental state perception, both of which are highly implicated in the Theory of Mind. As such, this result also opens fruitful new avenues for research on the properties of social gaze communication such as relative contributions of directionality vs. mental content, developmental trajectory and expressions in special groups, variations with individual differences, and/or the underlying neural mechanisms.

Data availability statement

The datasets generated and analyzed for this study can be found in the Open Science Framework directory, osf.io/3xcqk.

Ethics statement

The studies involving humans were approved by the McGill Univeristy's Research Ethics Board Office. 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

FM: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. FC: Writing – review & editing, Conceptualization, Methodology, Supervision. JR: Conceptualization, Formal analysis, Funding acquisition, Investigation, 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 research was supported by the Graduate and Postgraduate Fellowships from NSERC-CREATE, and grants from the Social Sciences and Humanities Council of Canada (SSHRC), Natural Science and Engineering Council of Canada (NSERC), and William Dawson Funds.

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

Apperly, I. A., and Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychol. Rev.* 116, 953–970. doi: 10.1037/a0016923

Baker, L. J., Levin, D. T., and Saylor, M. M. (2016). The extent of default visual perspective taking in complex layouts. *J. Exp. Psychol. Hum. Percept. Perform.* 42, 508–516. doi: 10.1037/xhp0000164

Baron-Cohen, S. (1991). Precursors to a theory of mind: understanding attention in others. *Nat. Theor. Mind* 1:1.

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

Capozzi, F., Beyan, C., Pierro, A., Koul, A., Murino, V., Livi, S., et al. (2019). Tracking the leader: gaze behavior in group interactions. *iScience* 16, 242–249. doi: 10.1016/j.isci.2019.05.035

Capozzi, F., and Ristic, J. (2018). How attention gates social interactions. *Ann. N. Y. Acad. Sci.* 1426, 179–198. doi: 10.1111/nyas.13854

Capozzi, F., and Ristic, J. (2020). AttentionANDmentalizing? Reframing a debate on social orienting of attention. *Vis. Cogn.* 28, 97–105. doi: 10.1080/13506285.2020.1725206

Cole, G. G., and Millett, A. C. (2019). The closing of the theory of mind: a critique of perspective-taking. *Psychon. Bull. Rev.* 26, 1787–1802. doi: 10.3758/s13423-019-01657-y

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

Del Sette, P., Bindemann, M., and Ferguson, H. J. (2022). Visual perspective-taking in complex natural scenes. *Q. J. Exp. Psychol.* 75, 1541–1551. doi: 10.1177/17470218211054474

Flavell, J. H., Flavell, E. R., Green, F. L., and Wilcox, S. A. (1981). The development of three spatial perspective-taking rules. *Child Dev.* 52:356. doi: 10.2307/1129250

Friesen, C. K., and Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychon. Bull. Rev.* 5, 490–495. doi: 10.3758/bf03208827

Frischen, A., Bayliss, A. P., and Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol. Bull.* 133, 694–724. doi: 10.1037/0033-2909.133.4.694

Happé, F., and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25. doi: 10.1007/s10803-005-0039-0

Jarrold, C., Butler, D. W., Cottington, E. M., and Jimenez, F. (2000). Linking theory of mind and central coherence bias in autism and in the general population. *Dev. Psychol.* 36, 126–138. doi: 10.1037//0012-1649.36.1.126

Kessler, K., and Rutherford, H. (2010). The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. *Front. Psychol.* 1, 108–119. doi: 10.3389/fpsyg.2010.00213

Kimhi, Y. (2014). Theory of mind abilities and deficits in autism spectrum disorders. *Top. Lang. Disord.* 34, 329–343. doi: 10.1097/tld.0000000000000033

Kingstone, A., Kachkovski, G., Vasilyev, D., Kuk, M., and Welsh, T. N. (2019). Mental attribution is not sufficient or necessary to trigger attentional orienting to gaze. *Cognition* 189, 35–40. doi: 10.1016/j.cognition.2019.03.010

Kuhn, G., and Kingstone, A. (2009). Look away! Eyes and arrows engage oculomotor responses automatically. *Attention, Perception, Psychophysics.* 71, 314–327. doi: 10.3758/APP.71.2.314

Lockwood, P. L., Apps, M. A. J., and Chang, S. W. C. (2020). Is there a "social" brain? Implementations and algorithms. *Trends Cogn. Sci.* 24, 802–813. doi: 10.1016/j.tics.2020.06.011

Marotta, A., Lupiáñez, J., Martella, D., and Casagrande, M. (2012). Eye gaze versus arrows as spatial cues: two qualitatively different modes of attentional selection. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 326–335. doi: 10.1037/a0023959

McCrackin, S. D., and Itier, R. J. (2019). Individual differences in the emotional modulation of gaze-cuing. *Cognit. Emot.* 33, 768–800. doi: 10.1080/02699931.2018.1495618

Moriguchi, Y., Ohnishi, T., Lane, R. D., Maeda, M., Mori, T., Nemoto, K., et al. (2006). Impaired self-awareness and theory of mind: an fmri study of mentalizing in alexithymia. *NeuroImage* 32, 1472–1482. doi: 10.1016/j.neuroimage.2006.04.186

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/s0140525x00076512

Qureshi, A. W., and Monk, R. L. (2018). Executive function underlies both perspective selection and calculation in level-1 visual perspective taking. *Psychon. Bull. Rev.* 25, 1526–1534. doi: 10.3758/s13423-018-1496-8

Ristic, J., Friesen, C. K., and Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Bull. Psychon. Soc.* 9, 507–513. doi: 10.3758/bf03196306

Ristic, J., and Kingstone, A. (2012). A new form of human spatial attention: automated symbolic orienting. *Vis. Cogn.* 20, 244–264. doi: 10.1080/13506285.2012.658101

Rosello, B., Berenguer, C., Baixauli, I., García, R., and Miranda, A. (2020). Theory of mind profiles in children with autism spectrum disorder: adaptive/social skills and pragmatic competence. *Front. Psychol.* 11, 17–30. doi: 10.3389/fpsyg.2020.567401

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., and Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1255–1266. doi: 10.1037/a0018729

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., and Heyes, C. (2014). Avatars and arrows: implicit mentalizing or domain-general processing? *J. Exp. Psychol. Hum. Percept. Perform.* 40, 929–937. doi: 10.1037/a0035175

Santiesteban, I., Kaur, S., Bird, G., and Catmur, C. (2017). Attentional processes, not implicit mentalizing, mediate performance in a perspective-taking task: evidence from stimulation of the temporoparietal junction. *NeuroImage* 155, 305–311. doi: 10.1016/j.neuroimage.2017.04.055

Sulpizio, V., Committeri, G., Metta, E., Lambrey, S., Berthoz, A., and Galati, G. (2015). Visuospatial transformations and personality: evidence of a relationship between visuospatial perspective taking and self-reported emotional empathy. *Exp. Brain Res.* 233, 2091–2102. doi: 10.1007/s00221-015-4280-2

Todd, A. R., Cameron, C. D., and Simpson, A. J. (2017). Dissociating processes underlying level-1 visual perspective taking in adults. *Cognition* 159, 97–101. doi: 10.1016/j.cognition.2016.11.010

Tomei, A., Besson, J., and Grivel, J. (2017). Linking empathy to visuospatial perspective-taking in gambling addiction. *Psychiatry Res.* 250, 177–184. doi: 10.1016/j.psychres.2016.12.061

Toscano, R., Price, G., and Scheepers, C. (2018). The impact of CEO arrogance on top management team attitudes. *Eur. Bus. Rev.* 30, 630–644. doi: 10.1108/ebr-12-2016-0156

Van Selst, M., and Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A* 47, 631–650. doi: 10.1080/14640749408401131

Wiese, E., Wykowska, A., Zwickel, J., and Müller, H. J. (2012). I see what you mean: how attentional selection is shaped by ascribing intentions to others. *PLoS One* 7:e45391. doi: 10.1371/journal.pone.0045391

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 24 October 2024
DOI 10.3389/fpsyg.2024.1458101

OPEN ACCESS

EDITED BY
Douglas Frye,
University of Pennsylvania, United States

REVIEWED BY
Antonio Narzisi,
Stella Maris Foundation (IRCCS), Italy
Ifigeneia Dosi,
Democritus University of Thrace, Greece

*CORRESPONDENCE
Meng-Jung Liu
✉ mj@mail.nknu.edu.tw

RECEIVED 01 July 2024
ACCEPTED 08 October 2024
PUBLISHED 24 October 2024

CITATION
Liu M-J (2024) Adolescents with autism spectrum disorder exhibit intact physical causal inference but weak intention inference.
Front. Psychol. 15:1458101.
doi: 10.3389/fpsyg.2024.1458101

Adolescents with autism spectrum disorder exhibit intact physical causal inference but weak intention inference

Meng-Jung Liu*

Department of Special Education, National Kaohsiung Normal University, Kaohsiung, Taiwan

Individuals with autism spectrum disorder (ASD) show impaired mentalizing skills, specifically in understanding intentions. They have difficulty understanding social situations with multiple cues due to their limited ability to perceive subtle social contextual cues. Studies that used comic strips and the strange stories as intention attribution tests found that individuals with ASD exhibit a reduced ability in attributing intentions compared to inferring causal consequences. This study aims to use static photographs of social scenes taken in everyday settings to investigate the ability of adolescents with ASD to infer intentions in social contexts, and to explore how intention inference relates to working memory and basic attention, including sustained attention, selective attention, and divided attention. The results show that the physical causal inference ability of adolescents with ASD is comparable to typically developing adolescents, whereas intention inference is notably weaker. Furthermore, working memory predicts physical causal inference and divided attention predicts intention inference in ASD.

KEYWORDS

autism spectrum disorder, theory of mind, mentalizing, intention attribution, physical causal inference, intention inference

Introduction

Autism spectrum disorder (ASD) is characterized by core deficits in social cognition and interaction. A review comparing social and non-social cognitive deficits in adults with ASD indicated that impairments are strongest in the mentalizing domain (Velikonja et al., 2019). Mentalizing is a form of social cognition to comprehend both one's own and others' behaviors in terms of internal states, such as thoughts, feelings, desires, and goals (Luyten et al., 2020). Studies suggest that adults with ASD struggle significantly with their ability to infer the mental states of others (Happé, 1994; Baron-Cohen et al., 2001; Castelli et al., 2002; Mazza et al., 2022). Understanding others' mental states (known as theory of mind) is a significant evolutionary achievement that deeply impacts our interactions with the world (Tomasello et al., 2005). Senju (2012) suggested two explanations for mentalization development failures in those with autism: one is that compensatory strategies may not be fully developed to allow for spontaneous and rapid mentalizing when needed, and the other is social cognition may not be automatically triggered by socially relevant cues.

Atypical social attention in ASD is widely discussed as a possible reason for difficulties in mentalizing. Social attention refers to the motivation, coordination of attention, and the focus on social cues within a contextual framework during interactions with others (Salley and Colombo, 2016). Individuals with ASD exhibit atypical attentional responses to social stimuli (Klin et al., 2002; Pelphrey et al., 2002; Corden et al., 2008; Norbury et al., 2009; Riby and Hancock, 2009a, 2009b; Nakano et al., 2010), including reduced attention to the eye and facial regions of others. In fact, when attending to faces, individuals with ASD tend to fixate more on the mouth area, and in situations presenting both social and non-social information concurrently, they may be more inclined to attend to body parts, objects, or stimuli in the background (Speer et al., 2007). Despite detecting similar information as neurotypical individuals, individuals with ASD demonstrate weaker sustained attention to social cues and a reduced ability to follow others' gazes in social contexts (Fletcher-Watson et al., 2009; Freeth et al., 2010). Swettenham et al. (1998) proposes that the attentional patterns seen in individuals with ASD are more indicative of general attentional abnormalities rather than a deliberate avoidance of social stimuli.

Furthermore, adolescents with ASD have weaker abilities than typically developing peers in perceiving subtle social contextual cues (Hodgins et al., 2020), making it difficult for them to comprehend social situations that involve multiple social cues. Pierce et al. (1997) used videos depicting social interactions with 1 to 4 cues, including verbal, tonal, non-verbal with objects, and non-verbal without objects, to assess social perception in children with ASD, finding that they perform similarly to typically developing children on general attention tasks such as counting characters and identifying genders. However, their ability to interpret social perception questions, especially in scenarios that include multiple cues, is notably weaker than that of typically developing children. Relatedly, Vivanti et al. (2011) observed significant impairments in adolescents with ASD regarding their ability to attend to and interpret referential cues, such as interpreting a head turn to understand someone's intentions. Contextual considerations alongside perceptual cues are crucial for judging interpersonal behaviors. Loveland et al. (2001) investigated how adolescents with ASD judge the appropriateness of social behaviors in videos containing multiple verbal and non-verbal cues. They found that adolescents with ASD are more likely to provide explanations of the protagonist's behavior that are irrelevant or idiosyncratic to the context, whereas typically developing adolescents are more likely to interpret behaviors in accordance with social norms and principles.

Studies show that theory of mind ability and working memory are moderately correlated (Davis and Pratt, 1995; Gordon and Olson, 1998; Hughes, 1998; Keenan et al., 1998; Carlson et al., 2002; Meyer and Collier, 2020; Imanipour et al., 2021). According to Ericsson and Kintsch (1995), working memory is a subset of information drawn from long-term memory, distinguished by contextual markers that highlight its relevance to ongoing cognitive processing. Some researchers propose that children performing the theory of mind (false-belief). task must inhibit incorrect responses based on factual knowledge stored in working memory (Carlson and Moses, 2001; Leslie et al., 2004). Others suggest that children need to switch their attention between their own mental state and others' mental states, both of which are managed in working memory (Frye et al., 1995; Andrews et al., 2003). Integrating both perspectives, with the ability to hold conflicting perspectives in mind, is essential for both learning and demonstrating knowledge related to theory of mind.

Social interactions rely on various interconnected processes such as attention, which is needed to interpret social cues and monitor others' actions and intentions (Capozzi and Ristic, 2018; Dosi and Boni, 2023). These processes work together dynamically to facilitate effective social engagement. McGlamery et al. (2007) have concluded that attention is predictive of theory of mind scores for kindergarten boys. Basic attention is crucial for daily functioning and cognitive tasks in perception and learning, laying the groundwork for more advanced cognitive processes such as memory encoding and problem-solving. Capozzi and Ristic (2018) identified three core processes that collaborate with attentional systems to influence selective responses to the social environment: perception, interpretation, and evaluation. Perception helps in prioritizing social cues, interpretation connects attention to understanding the social significance of cues and others' mental states, and evaluation assesses the value of social information sources. These processes collectively enable attention to effectively manage the vast amount of social information that one encounters daily.

Individuals with ASD show impaired mentalizing skills, specifically in understanding intentions, observed across both children and adults (Vivanti et al., 2011; Schneider et al., 2013; Schuwerk et al., 2016). They show reduced awareness of both their own and others' intentions, which is linked to broader impairments in theory of mind (Williams and Happé, 2010). Bodner et al. (2015) used stories to elicit responses that described physical relationships and mental or emotional states. The study found that individuals with ASD may lack stored experiential knowledge needed for specific inferences, possibly due to linguistic limitations that hinder access to relevant experiences. Baron-Cohen et al. (1986) and Le Donne et al. (2023) used non-verbal comic strips to investigate intention attribution and physical causal inference in children and adults with ASD. The current study aims to investigate the ability of adolescents with ASD to infer intentions in natural social contexts, while mitigating the constraints posed by their linguistic limitations. The test materials used in the current study were static photographs of social scenes taken in everyday settings. The use of naturalistic test formats and stimuli in studying subtle mind-reading deficits related to autism can potentially enhance task sensitivity because these measures are designed to mimic the challenges of everyday social interactions, which may uncover difficulties in the real-time processing of mental states (Heavey et al., 2000). Some studies used photographs of social scenes featuring real people as testing instruments (Riby and Hancock, 2008; Rice et al., 2012; Harrop et al., 2019; Ioannou et al., 2020), but these studies investigated the visual social attention or orientation of individuals with autism. To the author's knowledge, there have been no previous studies using photographs of natural social contexts to investigate characters' intention and physical causal inference in individuals with ASD. Using real interpersonal interaction scenarios as testing instruments will be closer to real-life situations, and the results obtained will more accurately reflect the daily life intention reasoning ability of those with autism.

The current study also explores how intention inference relates to working memory and basic attention, including sustained attention, selective attention, and divided attention. Sustained attention is the ability to maintain focus over extended periods. Selective attention refers to the cognitive process of concentrating on particular stimuli while filtering out irrelevant or unattended input, allowing for focused processing of selected information. Divided attention explores the challenges of multitasking and the optimal allocation of cognitive resources between tasks, emphasizing the need to split or quickly shift focus due to the inability to process all information simultaneously (Parasuraman, 1998).

Aims and research questions

The research questions of the current study explore the difference in intention reasoning abilities between adolescents with autism and typically developing adolescents, as well as the relationship between intention reasoning ability and working memory, sustained attention, selective attention, and divided attention.

Materials and methods

Participants

Pilot participants

The pilot participants consisted of 98 typically developing adolescents aged 11–18 years. Participants were recruited through school teachers who solicited willing participants from junior high and high schools. After distributing informed consent forms to students and parental consent forms, paper-and-pencil group tests were conducted at schools. Selection of schools took into account administrative district differences and gender ratios of participants.

Formal participants

The study included participants from an ASD group and a typically developing (TD) group. Recruitment for the ASD group was facilitated through autism foundations and their local branches nationwide, as well as by contacting schoolteachers to disseminate information. Inclusion criteria for the ASD group were as follows: (1) receiving a diagnosis of Asperger's syndrome or mild ASD by a psychiatrist; (2) ages between 11 and 18 years and attending regular classes; (3) Full Scale IQ above 85 measured by Wechsler Intelligence Scale for Children, Fourth Edition; (4) no comorbidities such as attention deficit hyperactivity disorder, epilepsy, anxiety disorder, or depression. Recruitment for the TD group prioritized students in the same class or school as the ASD participants. Teachers helped exclude students with other disabilities and to control for age and gender ratios between the two groups at their schools. Finally, there were 32 participants in the ASD group with a mean age of 15.6 years and 33 participants in the TD group with a mean age of 15.5 years, with 4 females in each group.

The ASD group and the TD group were matched by the schools they attended to control for urban–rural disparities. The results of the socioeconomic status survey showed no significant differences between the ASD group and the TD group. Participants of both groups were born in Taiwan, and their parents are all of local nationality, with Mandarin being the primary language used.

Procedure

The design of the test items involved referencing relevant research materials and collecting scenarios that align with the life experiences of typically developing adolescents aged 11–18 years. Real-life scenarios were captured using authentic photographs. Prior to the formal testing, the test items underwent content validity checks by two experts in the fields of psychology and special education. Subsequently, a pilot test was conducted using PowerPoint presentation software in a group setting, and based on the results, 4 items were removed. Each test item photo was displayed on a computer screen for 8 s.

The study was approved by the Institutional Review Board (IRB) of National Cheng Kung University in Taiwan. During the formal testing phase, both the TD group and the ASD group underwent individual testing. Participants were first given practice on 4 items, with the test administrator explaining the scenario depicted in each photo. The purpose of the practice was to ensure that participants understood the meanings of "accidently occurring" and "intentionally occurring." If participants made incorrect judgments, the test administrator would reiterate the scenario and the correct answer. Practice continued until participants responded correctly. Formal testing involved the test administrator asking participants, after viewing each photo, whether the depicted scenario was "accidently occurring" or "intentionally occurring," with the test administrator recording participants' verbal responses for each item.

Materials

Multi-dimension attention test

Multi-dimension attention test (Zhou et al., 1993) uses visual assessment materials to evaluate sustained attention, selective attention, and divided attention. Participants are required to circle corresponding items based on verbal instructions from the examiner under timed conditions. The test demonstrates a test–retest reliability coefficient ranging from 0.82 to 0.90 and an internal consistency coefficient ranging from 0.65 to 0.69.

Verbal working memory test

Verbal working memory test (Zeng, 1999) features common and high-frequency words from everyday life. Each item consists of 4 to 8 words. The examiner reads all words for each item aloud, provides an instruction, and asks participants to respond orally. For example, the examiner reads "desk, sausage, computer, dumpling, chicken leg and instructs, "Please verbally list the items that are edible in the original order." The test presents a total of 18 auditory items. Test–retest reliability ranges from 0.65 to 0.82.

Intention inferences test

The content of the test items was derived from social scenes in daily life. Each item was presented with a color photograph, retaining the environmental background and devoid of dialog. Each photo depicted one male and one female, both without overt facial expressions and avoiding direct eye contact with the camera to prevent direct eye contact with the participants and potential influence (von dem Hagen et al., 2014). The number of objects appearing in all photos was limited to six or fewer. The items were categorized into two types: physical causal inference (Figure 1) and intention inference (Figure 2). The difference between the two lies in inferring whether an event or behavior occurred accidently or was deliberately caused by someone. There were 26 items in total, with 13 items for each category. Each set of physical causal inference and intention inference items had identical scenes, differing only in the interaction between the characters.

The results of the factor analysis of the pilot indicated that, after orthogonal rotation using the minimum oblique rotation method, the former could explain 30.02% of the variance, and the latter could explain 8.62%, totaling 38.64%. There was a correlation of 0.621 between

FIGURE 1
Sample of physical causal inference.

FIGURE 2
Sample of intention inference.

the physical causal inference and intention inference factors, indicating a considerable relationship between the two factors. The Cronbach's α coefficient among items was 0.912, indicating good reliability.

Results

Table 1 presents descriptive statistics for physical causal inference and intention inference for ASD and TD groups. Table 2 shows the main effects of group differences reaching a significance level of 0.01, with a Partial Eta Squared of 0.199, indicating an effect size reaching the standard of Cohen's moderate effect size. The overall mean scores for both physical causal inference and intention inference were significantly higher in the TD group compared to the ASD group. The difference in inference types reached a significance level of 0.001, with a Partial Eta Squared of 0.393, indicating an effect size reaching the standard of Cohen's large effect size. Regardless of group, the mean score for physical causal inference was significantly higher than for intention inference. Furthermore, the interaction effect between inference types and groups reached a significance level of 0.001, with a Partial Eta Squared of 0.209, indicating an effect size reaching the standard of Cohen's moderate effect size. As the interaction effect between inference types and groups reached significance, it is not appropriate to directly interpret the test results for the main effects of factors between and within groups. Instead, further analysis is required to examine the simple main effects of factors between and within groups.

The results of the simple main effects ANOVA in Table 3 show that the difference between the two groups in physical causal inference did not reach a significance level, but the difference in intention inference did, with a Partial Eta Squared of 0.235, indicating an effect size reaching the standard of Cohen's moderate effect size. The results of the simple main effects ANOVA pairwise comparisons revealed that in physical causal inference, the difference in mean scores between the TD and ASD groups did not reach a significance level. However, in intention inference, the mean score for the TD group was significantly higher than that for the ASD group, reaching a significance level. The results of the simple main effects MANOVA for each type of inference in both ASD and TD groups show that the difference in both types of inference reached a significance level in the ASD group, with a Partial Eta Squared of 0.484, indicating an effect size reaching the standard of Cohen's large effect size. However, the difference in both types of inference did not reach a significance level of 0.05 in the TD group.

In order to investigate whether there are differences in the predictive accuracy of physical causal inference and intention inference between the TD and ASD groups regarding four variables: working memory, selective attention, divided attention, and sustained attention, the current study conducted a multiple group analysis of structural equation modeling (SEM) to compare the differences in regression coefficients between the two groups. As shown in Figures 3, 4, the results of the significance tests for regression coefficients revealed that none of the coefficients reached the 0.05 significance level for the TD group. However, in the ASD group, the regression coefficient for working memory predicting physical causal inference reached a significance level of 0.05 in a one-tailed test, the regression coefficient for divided attention predicting intention inference reached a significance level of 0.05 in a two-tailed test, and the regression coefficient for sustained attention predicting physical causal inference reached a significance level of 0.05 in a one-tailed test.

Regarding the test for differences in regression coefficients between the two groups, divided attention had a standardized regression coefficient of −0.21 in the TD group and as high as 0.33 in the ASD group. The test for the difference in regression coefficients between the two groups yielded a critical ratio (CR) of 2.097, with $p < 0.05$. This result indicates that the largest difference in predictive models between the two groups is in the ASD group, where the regression coefficient for divided attention predicting intention inference is significantly higher than the TD group.

The Specification Search method in the Amos software helps researchers choose a more concise and effective predictive model. In the current study, the Browne-Cudeck Criterion (BCC) was used to select the most parsimonious predictive model. The results, as shown in Figures 5, 6, indicate that for the TD group, the most parsimonious model retains only the correlations between predictor variables. The

TABLE 1 Descriptive statistics for physical causal inference and intention inference for ASD and TD groups.

Groups	*N*	Physical causal inference		Intention inference	
		Mean	SD	Mean	SD
ASD	32	12.47	0.63	10.37	1.83
TD	33	12.46	0.71	12.00	0.98

TABLE 2 Two-way mixed design ANOVA summary of groups and inference types.

Source	Type III sum of squares	df	Mean square	*F*	*p*	Partial Eta Squared
Between subjects						
Group (A)	18.46	1	18.46	13.42	0.001**	0.199
Ss w/in groups	74.31	63	1.18			
Within subjects						
Type of inference (B)	45.70	1	45.70	34.96	0.000***	0.393
A × B interaction	18.70	1	18.70	14.30	0.000***	0.209
B× Ss w/in groups	70.58	63	1.12			

***$p < 0.001$, **$p < 0.01$.

TABLE 3 Simple main effect ANOVA on physical causal inference and intention inference for ASD and TD groups.

Groups		SS	df	Mean Square	*F*	*p*	Partial Eta Squared	Compare means
Physical causal inference	Contrast	0.00	1	0.00	0.00	0.977	0.000	
	Error	23.93	63	0.38				
Intention inference	Contrast	37.16	1	37.16	16.50	0.000***	0.235	ASD < TD
	Error	120.97	63	1.92				

***$p < 0.001$.

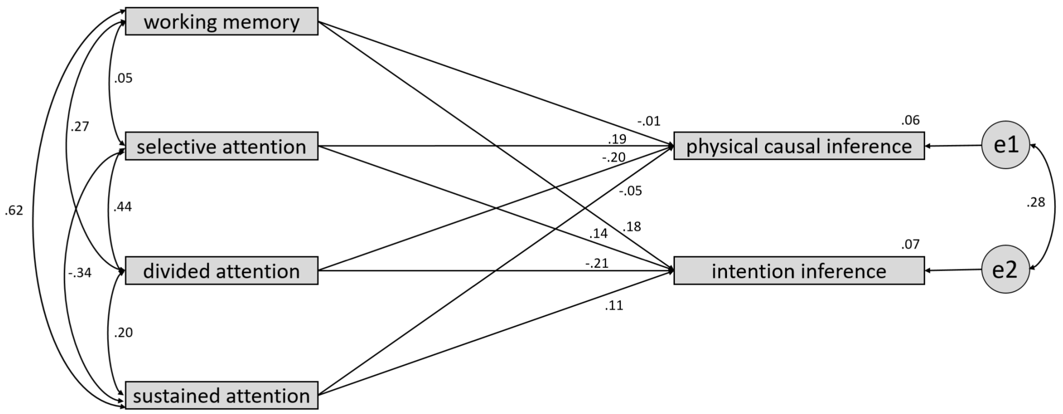


FIGURE 3
The regression coefficients between 4 variables and inference types for the TD group.

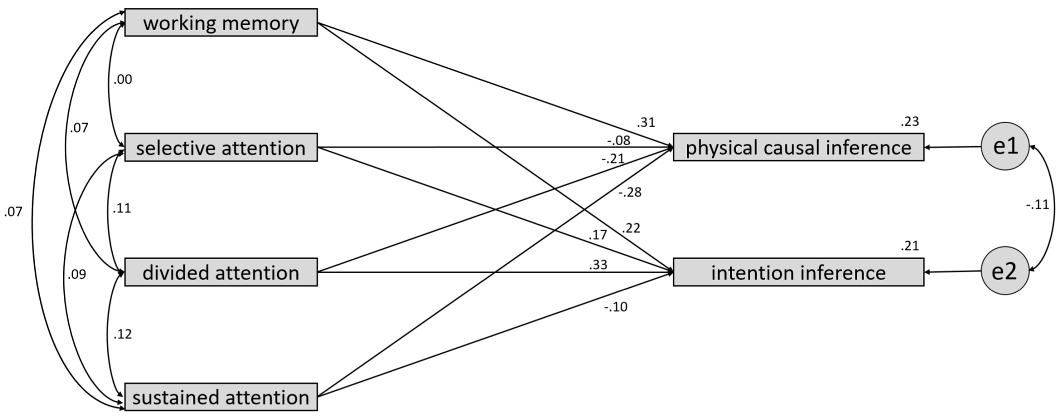


FIGURE 4
The regression coefficients between 4 variables and inference types for the ASD group.

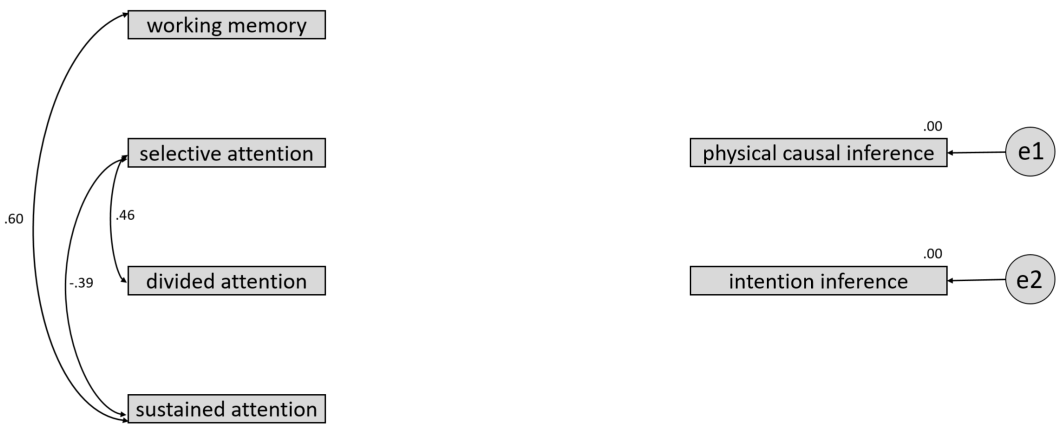


FIGURE 5
The most parsimonious model selected by the BCC for the TD group.

highest correlation observed is between working memory and sustained attention, reaching 0.60. The next highest correlation is between selective attention and divided attention, with a coefficient of 0.46. Notably, the correlation between selective attention and sustained attention is −0.39, suggesting that combining these two types of attention may not be suitable for the TD group.

The most parsimonious model for the ASD group is notably different from that of the TD group. According to the BCC, none of

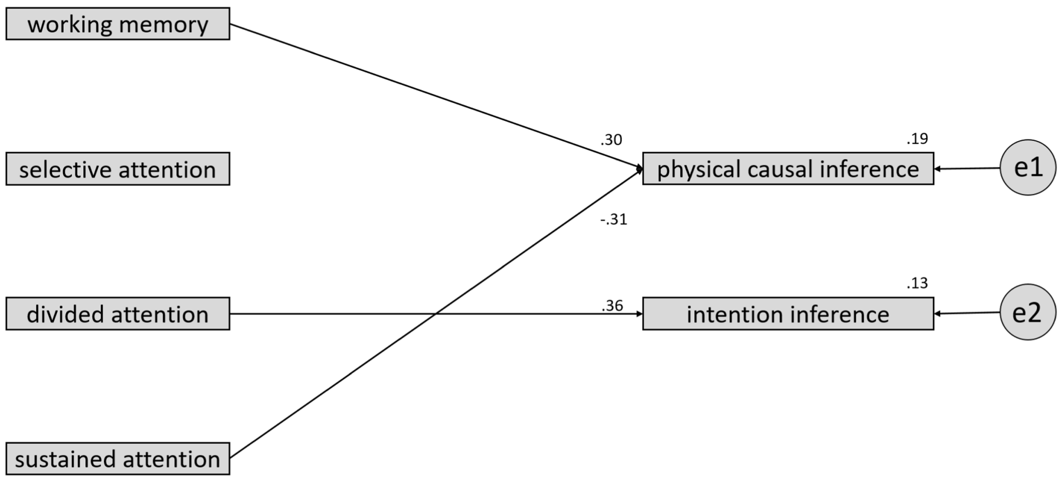


FIGURE 6
The most parsimonious model selected by the BCC for the ASD group.

the correlations between the 4 variables were retained. However, the standardized coefficients for working memory and sustained attention predicting physical causal inference were 0.30 and −0.31, respectively, and both were retained. Additionally, the standardized coefficient for divided attention predicting intention inference was as high as 0.36, making it the largest of the 3 coefficients retained. The R-squared values for working memory and sustained attention predicting physical causal inference were 0.19, while for divided attention predicting intention inference, it was 0.13, both reaching the standard of Cohen's moderate effect size.

Discussion

The results show that the physical causal inference of adolescents with ASD in the current study is comparable to typically developing adolescents, but the intention inference is notably weaker. The results support the belief that high-ability individuals with ASD can understand causal-mechanical conditions, but have difficulties in mentalizing, or the ability to understand and think about other people's thoughts and intentions (Baron-Cohen et al., 1986; David et al., 2008; Cole et al., 2018). Le Donne et al. (2023) used comic strips as an intention attribution test and found that the ASD group exhibited a reduced ability in attributing intentions compared to inferring causal consequences. In addition, studies using the strange stories tests have also shown consistent results in this direction (Happé, 1994; Jolliffe and Baron-Cohen, 1999; Kaland et al., 2005). As such, it appears that the use of different test content, including social scenes, comic strips, and the strange stories, have all produced the same result: compared to physical causal inference, individuals with ASD have remarkable difficulties in intention inference.

The social scenes used as test materials in the current study involve understanding the actions of the two characters in the photos. Understanding an action involves recognizing it on different levels: firstly, by identifying what was done and how; and secondly, by understanding why the action was performed and its effects (Kozak et al., 2006). Recognizing actions at the highest level suggests an awareness of one's own mind as the cause of behavior (Levy et al., 2002). This ability, known as action identification, enables tracking and interpreting mental states, applicable to both oneself and others. Boria et al. (2009) discovered that children with ASD fail to recognize an agent's actions on objects when those actions do not align with the standard use of the objects. Studies indicate that individuals with ASD process actions differently (Zalla et al., 2006; Vivanti et al., 2011; Kaiser and Pelphrey, 2012) and struggle with anticipating others' actions and representing goal-directed behaviors (Zalla et al., 2006; Chambon et al., 2017).

Another finding of this study is that divided attention predicts intention inference in ASD. Understanding other people's mental states involves a comprehensive analysis of their intentions, plans, personality, knowledge, emotions, beliefs, and desires. Additionally, it requires contextual understanding of the social situation in which events unfold (Killen et al., 2011). This holistic approach considers both individual psychological attributes and the broader social context as essential elements in understanding human behavior and interactions. The difficulties of people with autism are often reflected in terms of reduced attention to social cues in the environment (Leekam et al., 2000; Dawson et al., 2004; Shic et al., 2011). Our attentional system, typically involved in the pursuit of goal-directed behavior, serves the crucial function of selecting relevant stimuli and of ignoring irrelevant stimuli in different settings (Lavie et al., 2004). Frazier et al. (2017) suggest that in autism, failure to interpret people's actions might originate from basic atypicalities in selecting, differentially attending to, and/or integrating relevant information. The result of the current study shows divided attention predicts intention inference in adolescents with ASD. It may be a discrimination problem in selecting the most salient stimuli, sustaining attention to the most salient stimuli, or of filtering out extraneous information during visual perceptual experiences (Frazier et al., 2017). Divided attention is an executive function that involves the central executive component of working memory, allowing individuals to manage multiple tasks or sources of information simultaneously (D'Esposito et al., 1995). Fisher and Happé (2005) conducted a training study of theory of mind, and found autistic children who received the executive functioning training program performed comparably to those who received the theory of mind training program in the post-test. Future research could explore whether divided attention training or executive functioning training can enhance the intention reasoning abilities of individuals with autism.

In the current study, an unexpected finding is that physical causal inference is negatively correlated with sustained attention in adolescents with ASD. Most studies which examined the sustained attention of people with ASD reported no deficits (Garretson et al., 1990; Buchsbaum et al., 1992; Casey et al., 1993; Minshew et al., 1999; Noterdaeme et al., 2001; May et al., 2015), or conversely, have noted heightened attention in those with ASD, particularly on topics or objects that interest them (Plaisted and Davis, 2009). A possible explanation for the result is that the attention of adolescents with ASD may focus on irrelevant cues (within the visual field/among the visual stimuli) or attend to relevant cues without abstracting the accurate interpretation. Klin et al. (2002) conducted one of the first eye-tracking studies which revealed that individuals with ASD show altered patterns of social visual engagement when observing natural social settings, and that they exhibit a decreased focus on eyes, but an increased attention toward mouths, bodies, and objects. Similarly, Chita-Tegmark (2016) reviewed eye-tracking studies and concluded that when attending to social stimuli, individuals with ASD spend less time looking at the core features of the face, eyes, and mouth, but spend more time looking at bodies and attending to non-social elements. In addition, attenuated sensitivity to peripheral social targets is found in autistic children (Hou et al., 2024). Observations from visual tracking and verbal reports of individuals with ASD further indicate a diminished ability to perceive social cues crucial for contextual understanding (Tassini et al., 2022).

Studies have demonstrated that individuals with ASD exhibit atypical patterns of selective attention in both social and non-social contexts. Dawson et al. (1998) observed deviations in selective attention related to social information processing in ASD, indicating challenges in attending to and processing social cues typical for neurotypical individuals. Additionally, Renner et al. (2006) and Keehn et al. (2010) found that individuals with ASD show atypical selective attention in non-social information processing. This suggests differences in how individuals with ASD allocate attention to and process non-social stimuli compared to typically developing individuals. However, the current study did not find any relation between selective attention and physical causal inference or intention inference in ASD. This requires further investigation in future studies.

The result of this study reveals that working memory predicts physical causal inference in ASD, supporting the correlation between working memory and theory of mind. Lecce and Bianco (2018) investigated the role of working memory in theory of mind changes during middle childhood and indicated that individual differences in working memory moderated improvement in children's theory of mind. Accordingly, Lin et al. (2010) suggests that people with lower working memory capacity were less effective in applying their theory of mind to interpret behavior, and an attention-demanding task also reduced people's ability to apply their theory of mind. To study theory of mind understanding, Gregory et al. (2002) and Zalla et al. (2009) adjusted their studies by accounting for the impact of detailed attention on working memory. This adjustment aimed to ensure that working memory limitations did not confound their findings on theory of mind. Those studies confirm working memory and attention play crucial roles when applying theory of mind for typically developing individuals.

Limitations and future directions

The current study inevitably suffered from a range of limitations. This study used static photographs of social scenes to assess intention inference, which provided ecological validity. However, real-life social interactions are dynamic, and social cues appear rapidly, variables which cannot be accounted for in static photographs. It is suggested that future research consider using video clips of real social situations as testing materials to better approximate real-life contexts. This study did not test theory of mind or language abilities, so future research could analyze the relationship between theory of mind tests, language abilities, and intention reasoning. Since traits in female individuals with ASD may subtly differ from those in males (Hull et al., 2017), future research could include more females to investigate whether there are differences in intention reasoning abilities between genders in individuals with autism.

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 National Cheng Kung University Human Research Ethics Committee. 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. Written informed consent was obtained from the minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

Author contributions

M-JL: 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 work was supported by the National Science and Technology Council under Grant 105-2511-S-017-001.

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.

Publisher's note

References

Andrews, G., Halford, G. S., Bunch, K. M., Bowden, D., and Jones, T. (2003). Theory of mind and relational complexity. *Child Dev.* 74, 1476–1499. doi: 10.1111/1467-8624.00618

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *Br. J. Dev. Psychol.* 4, 113–125. doi: 10.1111/j.2044-835X.1986.tb01003.x

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). "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

Bodner, K. E., Engelhardt, C. R., Minshew, N. J., and Williams, D. L. (2015). Making inferences: comprehension of physical causality, intentionality, and emotions in discourse by high-functioning older children, adolescents, and adults with autism. *J. Autism Dev. Disord.* 45, 2721–2733. doi: 10.1007/s10803-015-2436-3

Boria, S., Fabbri-Destro, M., Cattaneo, L., Sparaci, L., Sinigaglia, C., Santelli, E., et al. (2009). Intention understanding in autism. *PLoS One* 4:e5596. doi: 10.1371/journal.pone.0005596

Buchsbaum, M. S., Siegel, B. V. Jr., Wu, J. C., Hazlett, E., Sicotte, N., Haier, R., et al. (1992). Brief report: attention performance in autism and regional brain metabolic rate assessed by positron emission tomography. *J. Autism Dev. Disord.* 22, 115–125. doi: 10.1007/BF01046407

Capozzi, F., and Ristic, J. (2018). How attention gates social interactions. *Ann. N. Y. Acad. Sci.* 1426, 179–198. doi: 10.1111/nyas.13854

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

Carlson, S. M., Moses, L. J., and Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant Child Dev.* 11, 73–92. doi: 10.1002/icd.298

Casey, B. J., Gordon, C. T., Mannheim, G. B., and Rumsey, J. M. (1993). Dysfunctional attention in autistic savants. *J. Clin. Exp. Neuropsychol.* 15, 933–946. doi: 10.1080/01688639308402609

Castelli, F., Frith, C., Happé, F., and Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125, 1839–1849. doi: 10.1093/brain/awf189

Chambon, V., Farrer, C., Pacherie, E., Jacquet, P. O., Leboyer, M., and Zalla, T. (2017). Reduced sensitivity to social priors during action prediction in adults with autism spectrum disorders. *Cognition* 160, 17–26. doi: 10.1016/j.cognition.2016.12.005

Chita-Tegmark, M. (2016). Social attention in ASD: a review and meta-analysis of eye-tracking studies. *Res. Dev. Disabil.* 48, 79–93. doi: 10.1016/j.ridd.2015.10.011

Cole, E. J., Slocombe, K. E., and Barraclough, N. E. (2018). Abilities to explicitly and implicitly infer intentions from actions in adults with autism spectrum disorder. *J. Autism Dev. Disord.* 48, 1712–1726. doi: 10.1007/s10803-017-3425-5

Corden, B., Chilvers, R., and Skuse, D. (2008). Avoidance of emotionally arousing stimuli predicts social-perceptual impairment in Asperger's syndrome. *Neuropsychologia* 46, 137–147. doi: 10.1016/j.neuropsychologia.2007.08.005

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

David, N., Gawronski, A., Santos, N. S., Huff, W., Lehnhardt, F. G., Newen, A., et al. (2008). Dissociation between key processes of social cognition in autism: impaired mentalizing but intact sense of agency. *J. Autism Dev. Disord.* 38, 593–605. doi: 10.1007/s10803-007-0425-x

Davis, H. L., and Pratt, C. (1995). The development of children's theory of mind: the working memory explanation. *Aust. J. Psychol.* 47, 25–31. doi: 10.1080/00049539508258765

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., and Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *J. Autism Dev. Disord.* 28, 479–485. doi: 10.1023/a:1026043926488

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., et al. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Dev. Psychol.* 40, 271–283. doi: 10.1037/0012-1649.40.2.271

Dosi, I., and Boni, N. (2023). "Pragmatic and language abilities in autism spectrum disorder: a case study approach" in Advances in psychology, mental health, and behavioral studies. IGI Global's online bookstore. eds. B. Christiansen and H. C. Chandan, 71–84.

Ericsson, K. A., and Kintsch, W. (1995). Longterm working memory. *Psychol. Rev.* 102, 211–245. doi: 10.1037/0033-295X.102.2.211

Fisher, N., and Happé, F. (2005). A training study of theory of mind and executive function in children with autistic spectrum disorders. *J. Autism Dev. Disord.* 35, 757–771. doi: 10.1007/s10803-005-0022-9

Fletcher-Watson, S., Leekam, S. R., Benson, V., Frank, M. C., and Findlay, J. M. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia* 47, 248–257. doi: 10.1016/j.neuropsychologia.2008.07.016

Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., et al. (2017). A meta-analysis of gaze differences to social and non-social information between individuals with and without autism. *J. Am. Acad. Child Adolesc. Psychiatry* 56, 546–555. doi: 10.1016/j.jaac.2017.05.005

Freeth, M., Chapman, P., Ropar, D., and Mitchell, P. (2010). Do gaze cues in complex scenes capture and direct the attention of high functioning adolescents with ASD? Evidence from eye-tracking. *J. Autism Dev. Disord.* 40, 534–547. doi: 10.1007/s10803-009-0893-2

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

Garretson, H. B., Fein, D., and Waterhouse, L. (1990). Sustained attention in children with autism. *J. Autism Dev. Disord.* 20, 101–114. doi: 10.1007/BF02206860

Gordon, A. C. L., and Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *J. Exp. Child Psychol.* 68, 70–83. doi: 10.1006/jecp.1997.2423

Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S., 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

Happé, F. (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

Harrop, C., Jones, D., Zheng, S., Nowell, S., Schultz, R., and Parish-Morris, J. (2019). Visual attention to faces in children with autism spectrum disorder: are there sex differences? *Mol. Autism.* 10:28. doi: 10.1186/s13229-019-0276-2

Heavey, L., Phillips, W., Baron-Cohen, S., and Rutter, M. (2000). The awkward moments test: a naturalistic measure of social understanding in autism. *J. Autism Dev. Disord.* 30, 225–236. doi: 10.1023/A:1005544518785

Hodgins, Z., Kelley, E., Kloosterman, P., Hall, L., Hudson, C. C., Furlano, R., et al. (2020). Brief report: do you see what I see? The perception of bullying in male adolescents with autism spectrum disorder. *J. Autism Dev. Disord.* 50, 1822–1826. doi: 10.1007/s10803-018-3739-y

Hou, W., Zhao, W., and Li, J. (2024). Intact gesture cueing of attention but attenuated sensitivity to peripheral social targets in autistic children: an eye-tracking and pupillometric study. *Biol. Psychol.* 191:108822. doi: 10.1016/j.biopsycho.2024.108822

Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Dev. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Hull, L., Mandy, W., and Petrides, K. V. (2017). Behavioural and cognitive sex/gender differences in autism spectrum condition and typically developing males and females. *Autism* 21, 706–727. doi: 10.1177/1362361316669087

Imanipour, S., Sheikh, M., Shayestefar, M., and Baloochnejad, T. (2021). Deficits in working memory and theory of mind may underlie difficulties in social perception of children with ADHD. *Neurol. Res. Int.* 2021, 1–7. doi: 10.1155/2021/3793750

Ioannou, C., Seernani, D., Stefanou, M. E., Riedel, A., Tebartz van Elst, L., Smyrnis, N., et al. (2020). Comorbidity matters: social visual attention in a comparative study of autism spectrum disorder, attention-deficit/hyperactivity disorder and their comorbidity. *Front. Psychol.* 11:545567. doi: 10.3389/fpsyt.2020.545567

Jolliffe, T., and Baron-Cohen, S. (1999). The strange stories test: a replication with high-functioning adults with autism or Asperger syndrome. *J. Autism Dev. Disord.* 29, 395–404. doi: 10.1023/A:1023082928366

Kaiser, M. D., and Pelphrey, K. A. (2012). Disrupted action perception in autism: behavioral evidence, neuroendophenotypes, and diagnostic utility. *Dev. Cogn. Neurosci.* 2, 25–35. doi: 10.1016/j.dcn.2011.05.005

Kaland, N., Møller-Nielsen, A., Smith, L., Mortensen, E. L., Callesen, K., and Gottlieb, D. (2005). The strange stories test: a replication study of children and adolescents with Asperger syndrome. *Eur. Child Adolesc. Psychiatry* 14, 73–82. doi: 10.1007/s00787-005-0434-2

Keehn, B., Lincoln, A. J., Muller, R., and Townsend, J. (2010). Attentional networks in children and adolescents with autism spectrum disorder. *J. Child Psychol. Psychiatry* 51, 1251–1259. doi: 10.1111/j.1469-7610.2010.02257.x

Keenan, T., Olson, D. R., and Marini, Z. (1998). Working memory and children's developing understanding of mind. *Aust. J. Psychol.* 50, 76–82. doi: 10.1080/00049539808257537

Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., and Woodward, A. (2011). The accidental transgressor: morally-relevant theory of mind. *Cognition* 119, 197–215. doi: 10.1016/j.cognition.2011.01.006

Klin, A., Jones, W., Schultz, R. T., Volkmar, F., and Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch. Gen. Psychiatry* 59, 809–816. doi: 10.1001/archpsyc.59.9.809

Kozak, M. N., Marsh, A. A., and Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *J. Pers. Soc. Psychol.* 90, 543–555. doi: 10.1037/0022-3514.90.4.543

Lavie, N., Hirst, A., de Fockert, J. W., and Viding, E. (2004). Load theory of selective attention and cognitive control. *J. Exp. Psychol. Gen.* 133, 339–354. doi: 10.1037/0096-3445.133.3.339

Le Donne, I., Attanasio, M., Bologna, A., Vagnetti, R., Masedu, F., Valenti, M., et al. (2023). Autism and intention attribution test: a non-verbal evaluation with comic strips. *Ann. General Psychiatry* 22:29. doi: 10.1186/s12991-023-00461-2

Lecce, S., and Bianco, F. (2018). Working memory predicts changes in children's theory of mind during middle childhood: a training study. *Cogn. Dev.* 47, 71–81. doi: 10.1016/j.cogdev.2018.04.002

Leekam, S. R., López, B., and Moore, C. (2000). Attention and joint attention in preschool children with autism. *Dev. Psychol.* 36, 261–273. doi: 10.1037//0012-1649.36.2.261

Leslie, A. M., Friedman, O., and German, T. P. (2004). Core mechanisms in "theory of mind". *Trends Cogn. Sci.* 8, 528–533. doi: 10.1016/j.tics.2004.10.001

Levy, S. R., Freitas, A. L., and Salovey, P. (2002). Construing action abstractly and blurring social distinctions: implications for perceiving homogeneity among, but also empathizing with and helping others. *J. Pers. Soc. Psychol.* 83, 1224–1238. doi: 10.1037/0022-3514.83.5.1224

Lin, S., Keysar, B., and Epley, N. (2010). Reflexively mindblind: using theory of mind to interpret behavior requires effortful attention. *J. Exp. Soc. Psychol.* 46, 551–556. doi: 10.1016/j.jesp.2009.12.019

Loveland, K. A., Pearson, D. A., Tunali-Kotoski, B., Ortegon, J., and Cullen Gibbs, M. (2001). Judgment of social appropriateness by children and adolescents with autism. *J. Autism Dev. Disord.* 31, 367–376. doi: 10.1023/a:1010608518060

Luyten, P., Campbell, C., Allison, E., and Fonagy, P. (2020). The mentalizing approach to psychopathology: state of the art and future directions. *Annu. Rev. Clin. Psychol.* 16, 297–325. doi: 10.1146/annurev-clinpsy-071919-015355

May, T., Rinehart, N., Wilding, J., and Cornish, K. (2015). Attention and basic literacy and numeracy in children with autism spectrum disorder: a one-year follow-up study. *Res. Autism Spectr. Disord.* 9, 193–201. doi: 10.1016/j.rasd.2014.10.010

Mazza, M., Pino, M. C., Keller, R., Vagnetti, R., Attanasio, M., Filocamo, A., et al. (2022). Qualitative differences in attribution of mental states to other people in autism and schizophrenia: what are the tools for differential diagnosis? *J. Autism Dev. Disord.* 52, 1283–1298. doi: 10.1007/s10803-021-05035-3

McGlamery, M. E., Ball, S. E., Henley, T. B., and Besozzi, M. (2007). Theory of mind, attention, and executive function in kindergarten boys. *Emot. Behav. Difficult.* 12, 29–47. doi: 10.1080/13632750601135899

Meyer, M. L., and Collier, E. (2020). Theory of minds: managing mental state inferences in working memory is associated with the dorsomedial subsystem of the default network and social integration. *Soc. Cogn. Affect. Neurosci.* 15, 63–73. doi: 10.1093/scan/nsaa022

Minshew, N. J., Luna, B., and Sweeney, J. A. (1999). Oculomotor evidence for neocortical systems but not cerebellar dysfunction in autism. *Neurology* 52, 917–922. doi: 10.1212/wnl.52.5.917

Nakano, T., Tanaka, K., Endo, Y., Yamane, Y., Yamamoto, T., Nakano, Y., et al. (2010). Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. *Proc. R. Soc. Lond.* 277, 2935–2943. doi: 10.1098/rspb.2010.0587

Norbury, C. F., Brock, J., Cragg, L., Einav, S., Griffiths, H., and Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders. *J. Child Psychol. Psychiatry* 50, 834–842. doi: 10.1111/j.1469-7610.2009.02073.x

Noterdaeme, M., Amorosa, H., Mildenberger, K., Sitter, S., and Minow, F. (2001). Evaluation of attention problems in children with autism and children with a specific language disorder. *Eur. Child Adolesc. Psychiatry* 10, 58–66. doi: 10.1007/s007870170048

Parasuraman, R. (1998). "The attentive brain: issues and prospects" in The attentive brain. ed. R. Parasuraman (Cambridge, MA: MIT Press), 3–15.

Pelphrey, K. A., Sasson, N. J., Reznick, J., Paul, G., Goldman, B. D., and Piven, J. (2002). Visual scanning of faces in autism. *J. Autism Dev. Disord.* 32, 249–261. doi: 10.1023/A:1016374617369

Pierce, K., Glad, K. S., and Schreibman, L. (1997). Social perception in children with autism: an attentional deficit? *J. Autism Dev. Disord.* 27, 265–282. doi: 10.1023/a:1025898314332

Plaisted, G. K., and Davis, G. (2009). Perception and apperception in autism: rejecting the inverse assumption. Philos. *Trans. R. Soc. Lond. B Biol. Sci.* 364, 1393–1398. doi: 10.1098/rstb.2009.0001

Renner, P., Klinger, L. G., and Klinger, M. R. (2006). Exogenous and endogenous attention orienting in autism spectrum disorders. *Child Neuropsychol.* 12, 361–382. doi: 10.1080/09297040600770753

Riby, D. M., and Hancock, P. J. B. (2008). Viewing it differently: social scene perception in Williams syndrome and autism. *Neuropsychologia* 46, 2855–2860. doi: 10.1016/j.neuropsychologia.2008.05.003

Riby, D. M., and Hancock, P. J. B. (2009a). Looking at movies and cartoons: eye tacking evidence from Williams syndrome and autism. *J. Intellect. Disabil. Res.* 53, 169–181. doi: 10.1111/j.1365-2788.2008.01142.x

Riby, D. M., and Hancock, P. J. B. (2009b). Do faces capture the attention of individuals with Williams syndrome or autism? Evidence from tracking eye movements. *J. Autism Dev. Disord.* 39, 421–431. doi: 10.1007/s10803-008-0641-z

Rice, K., Moriuchi, J. M., Jones, W., and Klin, A. (2012). Parsing heterogeneity in autism spectrum disorders: visual scanning of dynamic social scenes in school-aged children. *J. Am. Acad. Child Adolesc. Psychiatry* 51, 238–248. doi: 10.1016/j.jaac.2011.12.017

Salley, B., and Colombo, J. (2016). Conceptualizing social attention in developmental research. *Soc. Dev.* 25, 687–703. doi: 10.1111/sode.12174

Schneider, D., Slaughter, V. P., Bayliss, A. P., and Dux, P. E. (2013). A temporally sustained implicit theory of mind deficit in autism spectrum disorders. *Cognition* 129, 410–417. doi: 10.1016/j.cognition.2013.08.004

Schuwerk, T., Jarvers, I., Vuori, M., and Sodian, B. (2016). Implicit mentalizing persists beyond early childhood and is profoundly impaired in children with autism spectrum condition. *Front. Psychol.* 7:1696. doi: 10.3389/fpsyg.2016.01696

Senju, A. (2012). Spontaneous theory of mind and its absence in autism spectrum disorders. *Neuroscientist* 18, 108–113. doi: 10.1177/1073858410397208

Shic, F., Bradshaw, J., Klin, A., Scassellati, B., and Chawarska, K. (2011). Limited activity monitoring in toddlers with autism spectrum disorder. *Brain Res.* 1380, 246–254. doi: 10.1016/j.brainres.2010.11.074

Speer, L., Cook, A. E., McMahon, W. M., and Clark, E. (2007). Face processing in children with autism: effects of stimulus contents and type. *Autism* 11, 265–277. doi: 10.1177/1362361307076925

Swettenham, J., Baron-Cohen, S., Charman, T., Cox, A., Baird, G., Drew, A., et al. (1998). The frequency and distribution of spontaneous attention shifts between social and non-social stimuli in autistic, typically developing, and non-autistic developmentally delayed infants. *J. Child Psychol. Psychiatry* 39, 747–753. doi: 10.1111/1469-7610.00373

Tassini, S. C. V., Melo, M. C., Bueno, O. F. A., and de Mello, C. B. (2022). Weak central coherence in adults with ASD: evidence from eye-tracking and thematic content analysis of social scenes. *Appl. Neuropsychol. Adult* 31, 657–668. doi: 10.1080/23279095.2022.2060105

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691. doi: 10.1017/S0140525X05000129

Velikonja, T., Fett, A. K., and Velthorst, E. (2019). Patterns of non-social and social cognitive functioning in adults with autism spectrum disorder: a systematic review and meta-analysis. *JAMA Psychiat.* 76, 135–151. doi: 10.1001/jamapsychiatry.2018.3645

Vivanti, G., McCormick, C., Young, G. S., Abucayan, F., Hatt, N., Nadig, A., et al. (2011). Intact and impaired mechanisms of action understanding in autism. *Dev. Psychol.* 47, 841–856. doi: 10.1037/a0023105

von dem Hagen, E. A. H., Stoyanova, R. S., Rowe, J. B., Baron-Cohen, S., and Calder, A. J. (2014). Direct gaze elicits atypical activation of the theory-of-mind network in autism spectrum conditions. *Cereb. Cortex* 24, 1485–1492. doi: 10.1093/cercor/bht003

Williams, D., and Happé, F. (2010). Representing intentions in self and other: studies of autism and typical development. *Dev. Sci.* 13, 307–319. doi: 10.1111/j.1467-7687.2009.00885.x

Zalla, T., Labruyere, N., and Georgieff, N. (2006). Goal-directed action representation in autism. *J. Autism Dev. Disord.* 36, 527–540. doi: 10.1007/s10803-006-0092-3

Zalla, T., Sav, A. M., Stopin, A., Ahade, S., and Leboyer, M. (2009). Faux pas detection and intentional action in Asperger syndrome: a replication on a French sample. *J. Autism Dev. Disord.* 39, 373–382. doi: 10.1007/s10803-008-0634-y

Zeng, S. J. (1999). Verbal working memory test. (Taipei, Taiwan: National Science and Technology Council).

Zhou, T. J., Qiu, S. Z., and Song, S. H. (1993). Multi-dimension attention test. (Taipei, Taiwan: Psychological Publishing).

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 02 December 2024
DOI 10.3389/fpsyg.2024.1435685

Check for updates

OPEN ACCESS

EDITED BY
Douglas Frye,
University of Pennsylvania, United States

REVIEWED BY
Christopher Osterhaus,
University of Vechta, Germany
Laura Visu-Petra,
Babeş-Bolyai University, Romania

*CORRESPONDENCE
Derek E. Montgomery
✉ montg@fsmail.bradley.edu

RECEIVED 20 May 2024
ACCEPTED 08 November 2024
PUBLISHED 02 December 2024

CITATION
Montgomery DE, Tompkins V and Feng X (2024) The theory of mind construct in adulthood: perspective taking in relation to language and executive function. *Front. Psychol.* 15:1435685. doi: 10.3389/fpsyg.2024.1435685

The theory of mind construct in adulthood: perspective taking in relation to language and executive function

Derek E. Montgomery[1]*, Virginia Tompkins[2] and Xin Feng[3]

[1]Department of Psychology, Bradley University, Peoria, IL, United States, [2]Department of Psychology, The Ohio State University at Lima, Lima, OH, United States, [3]Department of Human Sciences, The Ohio State University, Columbus, OH, United States

There are conflicting proposals about the underlying structure of the theory of mind (ToM) construct. The lack of clarity impedes attempts to understand relationships between ToM and other cognitive abilities. This study investigated the nature of the ToM construct and its relation to cognitive variables by administering a battery of ToM measurements along with measurements of executive function and general vocabulary to 207 (M_{age} = 19.26) adult participants. Associations between ToM tasks were statistically significant after controlling for covariates, but, for the most part, very weak in magnitude. The strongest relationship was between the Strange Stories and Higher-Order False Belief measurements. Previous theoretical analysis proposes those instruments are conceptually linked by a perspective taking requirement that entails representing another's mental state. Results from a factor analysis suggested an underlying ToM structure—a *protagonist perspective* factor. The Strange Stories, Higher-Order False Belief, and Frith-Happé Animation tasks loaded onto the factor. Its defining feature is the ascription of mental states to predict and explain protagonists' actions that take place within a narrative structure. It is related more strongly to vocabulary than executive function and it provides grounds for future research on the role of narrative processing in ToM reasoning.

KEYWORDS

theory of mind, individual differences, perspective taking, early adulthood, theory and measurement

1 Introduction

Theory of mind (ToM) is the attribution of cognitive, volitional, and affective states to predict and explain human action. Events that might otherwise appear isolated and incomprehensible are instead integrated within a causal-explanatory network of mental state concepts:

> *One of the prime functions of everyday belief-desire psychology is to make sense of human actions and minds, to fill in the "becauses" of human life, to answer the everyday "why" questions about what we are doing….* (Wellman, 2014, p. 42).

Research in early childhood focuses on mastery of basic mental state concepts like belief, knowledge, and desire. In story-like scenarios, preschoolers predict and explain what a character will do, or how a character will feel, based upon the character's mental states (Wellman, 2014). A hallmark of ToM development emerges around age 4 when children master false belief tasks (Wellman et al., 2001). False belief understanding is an index of the capacity to distinguish between an objective state of affairs and a subjective viewpoint about those affairs. Advanced ToM (after age

5) builds upon basic concepts and involves reasoning about misinterpretations and subtle motivations, and making distinctions between what people literally say and what they actually mean (Miller, 2022). Advanced ToM measurements are designed to capture competency in making sense of what someone says or does by integrating context (e.g., situations, personal histories), nonverbal cues, and mental states.

Knowing the adult endpoint of ToM informs research questions for studying its development. For instance, if a mature ToM is fractionated into different subcomponents, then attendant questions ask if the same differentiated structure is present in early ToM and, if so, whether there are different timetables and different underlying cognitive processes for each component. Clarity on the adult endpoint informs hypotheses about cognitive mechanisms underlying its development. There are parallel questions in related areas of developmental psychology. Research on the development of the executive functions (EF) has consistently identified a relatively undifferentiated beginning state in early childhood followed by differentiation into early adulthood (e.g., Karr et al., 2022). There is a rich empirical literature on the adult structure of the EF construct (e.g., Miyake and Friedman, 2012); in contrast, however, there is uncertainty about the structure of ToM in adulthood. We discuss different, conflicting, theories about the structure of ToM in the next section.

1.1 The theory of mind construct

There is a growing consensus in the field that the ToM construct, as it is typically defined, is too vague (e.g., Schurz et al., 2021). A three-step process driven by conceptual analysis and research is needed to clarify the construct, its assessment, and, ultimately, the processes that underlie its development (Happé et al., 2017). In the first step, theoretical analysis aims to clarify the construct by "deconstructing" it into criterial subcomponents (e.g., Quesque and Rossetti, 2020; Schaafsma et al., 2015; Schurz et al., 2021). The assumption behind deconstruction is that ToM is not a monolithic construct but, instead, there are varieties of ToM.

Under that assumption, the next steps involve aligning measurements with subcomponents of the construct and, lastly, identifying the neurocognitive processes that underlie task performance.

Taken as a whole, those steps line up the construct, its measurement, and the basic processes driving performance on ToM measurements. In this paper we attempt to advance that research agenda with a focus on the second step. We ask whether widely used ToM instruments cohere in a manner that is consistent with hypothesized subcomponents of the ToM construct. The focus on task coherence naturally leads to considerations about specific cognitive abilities that are common to the tasks that assess a particular subcomponent. We also address, as an attendant consideration in the paper, the relationship between general cognitive abilities and ToM.

The focus of this paper is on the mature form of ToM exhibited by the time individuals reach early adulthood. Although there is considerable research on the coherence of ToM tasks in early to middle childhood (e.g., Osterhaus et al., 2016), comparatively little is known about the coherence of such tasks in adulthood. Extrapolating results from childhood is complicated by the greater variety of task demands for measures of advanced ToM. Early childhood ToM tasks are generally brief and limited to pass-fail response options because of children's performance limitations (e.g., vocabulary, attention). In contrast, advanced ToM measurements have a much broader range of task demands (e.g., reading vignettes, completing self-report questions, articulating verbal responses to explain characters' actions, and interpreting ambiguous or subtle perceptual cues). Vast differences in task demands threaten to mask actual coherence among advanced ToM measurements (Devine, 2021). Therefore, advanced ToM tasks represent a stringent test of whether the ToM construct is unitary or coheres into meaningful subcomponents.

The tasks we evaluate are listed in Table 1. We chose them for three reasons. First, they are among the most frequently used instruments in research on advanced ToM (Osterhaus and Bosacki, 2022) with the exception of the newly developed Four-Item Mentalizing Index (FIMI) self-report measurement (Clutterbuck et al., 2021). We added that task to our battery because it targets a competency—cognitive perspective taking—that features prominently in an analysis of the ToM construct discussed below (Quesque and Rossetti, 2020). Second, the tasks in Table 1 (except for the FIMI) have

TABLE 1 Advanced theory of mind measurements.

Higher-order false belief Higher-order tasks assess the understanding that characters in scenarios can have false beliefs about beliefs (e.g., Perner and Wimmer, 1985). Higher-order tasks recursively embed desires and beliefs such as A wanted B to know that A believes… Questions range in complexity, with up to seven levels of embeddedness. The participant's perspective counts as the first level; thus, level-two embeddedness is about one mental state (e.g., "A wanted X"), level-three would be, "A thought that B wanted X," and so forth (Launay et al., 2015).
Reading-the-mind-in-the-eyes test (RMET) Photographs of the eye region depict expressions of emotion and thought. Participants choose a mental state descriptor (e.g., *suspicious*) from among four options that best describes what the person is thinking or feeling (Baron-Cohen et al., 2001; Olderbak et al., 2015).
Frith-Happé animation A brief video displays geometric shapes moving in a manner that suggests both contingency and narrative elements such as goals and emotional reactions. Participants' descriptions of what happened in the scenes are coded for mention of thoughts, desires, emotions, and other ToM-related constructs (Abell et al., 2000).
Strange stories A series of brief scenarios portray social interactions that feature a non-literal verbal utterance (e.g., a deceptive or ironic statement). Participants are asked to interpret the utterance and they are scored on whether they infer the speaker's intended meaning underlying the literal content of the utterance (White et al., 2009).
Four-item mentalizing index (FIMI) A self-report measurement of the ability to infer the cognitive perspective of others (Clutterbuck et al., 2021).

been administered across the lifespan, which means that results from this study are relevant to questions about whether there is developmental continuity in the ToM construct (Apperly et al., 2009). Related, there is variability in performance among adults on the instruments used in this study (e.g., Clutterbuck et al., 2021; Klindt et al., 2017; Launay et al., 2015; Murray et al., 2017; Osterhaus and Bosacki, 2022). Third, the tasks were chosen because they speak to a theoretical point that we elaborate upon in section 1.1.2; namely, that cognitive perspective taking is a defining component of ToM.

In order to gain clarity on the structure of the construct and the processes that support it, we address three questions in this paper:

a. Are there interrelations among some (or all) of the ToM measurements in Table 1?
b. If there are interrelations, do they cohere into theoretically meaningful clusters (subcomponents) or are the relations sufficiently widespread to suggest ToM is a unitary construct?
c. If there are interrelations, are they related to general cognitive (including language) skills and/or socioeconomic status?

Below, we expand upon questions a, b and c.

1.1.1 Are there interrelations among ToM measurements?

In principle, it is reasonable to expect that, all else equal, measurements of ToM might show at least some degree of interrelatedness despite differences in the tasks' demands, formats, and modalities. This is because mental state concepts of belief, desire, intention, perception, and emotion cohere to form a causal-explanatory network within the ToM construct (Wellman, 2014).

To illustrate, in false belief tasks an agent sees the location of an object but does not witness its transfer to another location (e.g., Wimmer and Perner, 1983). Participants infer where the agent will search for the object upon returning to look for it. Within that brief scenario, there is a complex interplay among the concepts of perception, belief, desire/volition, and emotion. The agent's desire to find the object motivates the search for it; however, the nature of the search (where the agent looks) is mediated by the belief about its location. The belief about its location is, in turn, tied to the agent's earlier perceptual experience. The resultant surprise and disappointment upon looking in the wrong location are tied to the agent's desire and false belief. Thus, while task content and demands will vary, "solutions" to the tasks will draw from the same causal-explanatory framework to make sense of the agent's actions by ascribing interrelated mental state concepts. In that sense, ToM is a unitary construct, leading one to expect a widespread degree of relatedness between diverse measures of ToM (see Warnell and Redcay, 2019 for a discussion).

At the other end of the continuum, Warnell and Redcay (2019) found that ToM tasks were unrelated to one another in adulthood (effect sizes ranged from $r = -0.115$ to $r = 0.125$). Their results suggest that ToM has a diverse, and perhaps even fragmented, structure. While instructive, there are limitations to the Warnell and Redcay study. First, their task choice prioritized identifying a diverse array of measurements, but it was not theoretically driven. They acknowledged in the discussion of their findings that, "future studies should employ a targeted set of tasks in order to test specific underlying structures of social cognition" (p. 8). Second, they did not include measurements of EF, general language skills, or SES as covariates in their adult sample. The inclusion of those constructs is important for both theoretical and methodological reasons, as discussed in section 1.1.3.

1.1.2 Perspective taking and theory of mind

The tasks in our study are different versions or differ entirely from those used in Warnell and Redcay (2019). With respect to task variants, both studies used RMET and the Higher-Order False Belief tasks, but the studies used different versions of the tasks (see section 2). The task choice in the present study is theoretically motivated, with a focus on a recent analysis of the ToM construct by Quesque and Rossetti (2020), which we discuss in this section. In particular, we ask if ToM measures cohere into a meaningful *perspective taking* component.

Meta-representation, the "conceptual heart" of ToM, involves (a) representing the content of others' mental states (i.e., their perspectives) and (b) doing so even when perspectives are in conflict between individuals and/or with reality (Rakoczy, 2022, p. 1). Quesque and Rossetti (2020) argued that those criteria are inconsistently assessed in tasks that purportedly measure ToM. In their analysis, some tasks, such as the Strange Stories and False Belief tasks, meet both criteria (Quesque and Rossetti, 2020). The FIMI, as a measure of individual differences in cognitive perspective taking, would presumably relate to both measurements. Other tasks, such as those that involve the recognition of mental states from perceptual stimuli, do not meet the criteria. The RMET and Animation tasks both fall into the latter category. Performance on those measurements, they argue, involves low level perceptual processes of recognition and categorization rather than actually representing the conflicting perspectives of different agents.

Meinhardt-Injac et al. (2020) found evidence for a two-component model of ToM in a study of participants between 11 and 25-years-old. The social-perceptual component entails inferring mental states from nonverbal cues (such as the eyes). Mental state judgments are immediate and directly elicited from perceptual stimuli. The social-cognitive component involves verbal reasoning about mental states. In distinguishing between social-perceptual and social-cognitive components, the two-factor model is consistent with Quesque and Rossetti (2020).

The Quesque and Rossetti (2020) analysis suggests that the Strange Stories and False Belief tasks (and the FIMI) should be related to one another, but not related to the Animation and RMET tasks. However, there is inconsistent support for that prediction within the childhood literature. On the one hand, measurements that should be related are, in fact, not always related. For instance, Hayward and Homer (2017) found that Second-Order False Belief and Strange Stories tasks were unrelated among 7–13-year-olds ($n = 107$; $r = 0.06$). Also, measurements that should be unrelated in the Quesque and Rossetti analysis sometimes actually are related. Devine (2021) suggests that the Animation and Strange Stories tasks share the same conceptual core. In both instances, participants explain social behavior by inferring the mental states of agents. In support of that characterization, the tasks were significantly correlated in 10-year-old children ($n = 137$; $r = 0.34$) and loaded onto a single latent factor (Devine et al., 2016; see also, Lecce et al., 2021).

Partial support comes from factor analytic work by Osterhaus et al. (2016). They found that three measurements—Higher-Order False Belief, Strange Stories, and the RMET—cohered into a social reasoning factor for children in grades 2–4 ($n = 466$). Two other factors were reasoning about ambiguity and recognition of social norm transgressions. The relationship between Higher-Order False Belief and Strange Stories is at odds with Hayward and Homer (2017). One explanation is task variation between studies. For instance, Hayward and Homer reported comprehension problems for the Strange Stories task in their sample. Consequently, they limited their analyses to just a subset of the items for which children passed comprehension questions. More generally, in the ToM literature there are multiple task variants, which complicates between-study comparisons and likely contributes to inconsistent findings.

Inconsistencies in the literature may also stem from different age ranges found between samples. There was a narrower age range in Osterhaus et al. (2016) than in Hayward and Homer (2017). In general, the narrower the age range, the less likelihood there is of conflating developmental and individual differences (see Devine, 2021). The problem with a large range is that variability in performance on a particular task might be minimal for children at the low or high age range of a sample. For instance, children at the low range of the sample may have a conceptual deficit that severely restricts performance relative to children at the upper age range of the sample. Alternatively, children in the upper age range may have reached ceiling on a task.

Although comparisons between studies are difficult for the reasons given above, the point that is most germane to the present study is simply that there is a need to empirically test the Quesque and Rossetti (2020) analysis because support is not unequivocal in the literature. The patchwork of findings highlight the need for research that will inform whether the Quesque and Rossetti analysis actually captures the mature form of the ToM construct. In particular, there is a need for research that tests ToM performance on a battery of widely used measures administered to a large adult sample that has a restricted age range. A narrow age range, for the reasons noted above, "is essential for studying individual differences in ToM and for examining the relations between tasks" (Devine, 2021, p. 61).

1.1.3 Cognitive influences

Studies of adult ToM that are focused on relationships with language and EF are uncommon in the literature (Miller, 2022). Even when studies of advanced ToM do include measures of EF and language, Miller points out that frequently the goal is to confirm equivalency between groups rather than investigate their influence on ToM. A further complication for understanding how ToM is related to language and EF in adulthood stems from the tendency to analyze ToM with a composite score. Doing so limits our understanding of how general cognitive influences might vary across different measures. Put another way, there is an incomplete picture of how language and EF relate to performance on a range of different measures of advanced ToM in adulthood.

Osterhaus and Bosacki (2022) found that general language skills (mostly vocabulary) and EF (inhibition) were related to individual differences in advanced ToM in their analysis of studies across the lifespan. They calculated the average correlation by collapsing across tasks and age groups. Both correlations were statistically significant, but the magnitude of the relation was greater between language and advanced ToM ($r = 0.376$) than for inhibition ($r = 0.152$). Because Osterhaus and Bosacki (2022) included a wide age range in their analyses, their findings do not speak directly to the potential influence of cognitive variables on advanced ToM in adulthood. Nevertheless, the results indicate a potentially consequential role for both variables and warrant inclusion of both language and EF as covariates in the present study.

Accounting for language and EF in the present study allows us to address a theoretical question surrounding their respective roles in adults' ToM. Language and EF are associated with ToM development in early childhood (e.g., Tompkins et al., 2019; Devine and Hughes, 2014). The variables are posited to promote both the emergence and expression of ToM in childhood (Carlson et al., 2013). *Expression* refers to task performance; for instance verbal tasks require general language skills for comprehending scenarios and test questions. *Emergence* refers to a role for language and EF in producing variation in the timing of conceptual development. For instance, preschoolers' understanding of complement syntax supports the development of false belief understanding (de Villiers, 2021). Language in that case supports the development of a core ToM concept.

Core concepts of belief and desire are in place prior to advanced ToM. Therefore, it is unlikely that variability in task performance on advanced ToM tasks would be due to conceptual deficiencies *per se*:

> "Thinking in terms of mindreading concepts does not help us understand variability of mindreading in typical adults, because there is no variance in the possession of such concepts beyond late childhood" with Apperly and Wang (2021, pp. 100–101).

Variability in advanced ToM reflects individual differences in the fluency of social reasoning and in sensitivity to subtle social cues rather than mastery of basic concepts. Advanced ToM tasks are not designed to capture conceptual deficits in an all-or nothing (pass-fail) manner; instead, the tasks generally measure the consistency with which conceptual knowledge is applied to problems in social interactions. From that perspective, it seems more likely that language and EF, if they are responsible for individual differences in advanced ToM, would have an impact on the expression of ToM reasoning instead of the development of new conceptual knowledge (Apperly, 2021).

The question, then, is whether the influence of language and EF in the development of conceptual knowledge in childhood endures in the expression of that knowledge in adulthood, or whether linguistic and EF skills have reached a level of proficiency by early adulthood that is sufficient for mastering advanced ToM tasks (Apperly, 2021). In the latter case, language and EF would no longer be limiting factors in task performance and, therefore, would be unrelated to individual differences in ToM. Evidence consistent with the latter possibility comes from a recent meta-analysis of individual differences in EF and second-order false belief understanding (Peloquin et al., 2023). The magnitude of the relationship between EF and second-order false belief decreased with age, and by adulthood it was non-significant. However, with one exception, the studies only employed a second-order false belief task, and that task is subject to ceiling effects in adulthood (Osterhaus and Bosacki, 2022). The false belief measurement in our study (Launay et al., 2015) includes higher-order questions that recursively embed mental states beyond the second order (see Table 1). As noted earlier, we deliberately chose that task because it produces variability in adult samples (Launay et al., 2015).

A third variable of interest is the socioeconomic status (SES) of the participants. SES is a common covariate in ToM research in childhood, and there is evidence that it is related to ToM development (e.g., Tompkins et al., 2017; see Devine and Hughes, 2018 for a meta-analysis). To be clear, SES itself is not a cognitive variable, but it is a proxy for a wide range of variables that include education and parental behaviors associated with children's language and cognitive development. Longitudinal evidence indicates that SES differences in early childhood are associated with educational status and general cognitive skills in young adulthood (Osler et al., 2013). Consequently, it is possible the SES-ToM relationship observed in childhood may extend into early adulthood. However, there is relatively little research examining that possibility.

In sum, the first research question for this study asks if different measurements within a ToM battery are interrelated even when accounting for covariates. If the ToM construct is unified, then commonly used measurements of the construct should be interrelated. If the ToM construct has an unstable structure, then tasks should generally be unrelated because there is not a single, foundational core component common to all (e.g., Warnell and Redcay, 2019). The second question asks if factor analysis reveals evidence of coherence among the tasks into a theoretically meaningful factor or factors. The factor would represent a subcomponent of ToM. One candidate subcomponent is a "perspective taking" factor that would include Higher-order False Belief, Strange Stories, and the FIMI tasks (Quesque and Rossetti, 2020). Finally, if evidence indicates the presence of a factor, a third question asks how the factor is related to other cognitive variables. Individual differences in ToM in childhood are associated with vocabulary and EF. If the influence of those variables is continuous, then we would expect them to also bear a relation to individual differences in adulthood (see Apperly, 2021).

2 Methods

2.1 Participants

Undergraduate students were recruited from Introductory Psychology courses on two college campuses in the Midwest United States. The study received Institutional Review Board approval from The Ohio State University; written consent was obtained for all participants; and participants received course credit for their participation. Participants (76% female) included 207 adults ($M_{age} = 19.26$, $SD = 1.13$, range = 18.02–26.69); 49 identified as male, 158 as female, and no participants identified as non-binary or chose to self-describe. The sample size is sufficient for uncovering lowest practically meaningful correlation coefficients ($r = 0.20$). The race of participants was 74% White, 15% Black, 7% Asian, 3% Other, and < 1% American Indian or Alaska Native; assessed separately, 12% also identified as Hispanic or Latino. Race and ethnicity closely mirror proportions in the U.S. population (U.S. Census Bureau, 2022).

2.2 Measures

2.2.1 Theory of mind

Reading the mind in the eyes. This task presents four sets of eyes and participants choose the mental state that best matches the expression in the eyes. We used the shortened version of 10 sets of eyes (Olderbak et al., 2015) because it has better internal reliability than the original 36 sets of eyes (Baron-Cohen et al., 2001) used in the Warnell and Redcay (2019) study. Pages (10 plus one practice item) with eyes and four answer choices were presented in a binder and participants recorded their responses on an answer sheet. For example, for the target mental state "pensive," the other three answer choices were *irritated*, *excited*, and *hostile*. Definitions of all answer choices were provided in the binder. We calculated Kuder–Richardson 20 (KR20) for internal reliability for the eyes task given its binary scoring (0 or 1), which was 0.31.

Frith-Happé animation. The Animation task (Abell et al., 2000; White et al., 2011) presents participants with a big red triangle and a small blue triangle moving about a white background. Animations varied in length between 34 and 45 s and were presented on an iPad. We administered one practice animation and the four ToM animations (coaxing, mocking, seducing, surprising). To illustrate, in one of the animations scenarios a big triangle and a small triangle interact in a manner that suggests one is coaxing the other to come out of a box-shaped room. At the end of each animation, participants were asked "What was happening in this animation?." Responses were recorded and transcribed verbatim. Each response was coded for the level of intentionality according to Lecce et al. (2021). Scores ranged from 0 to 5 on each animation for a total of 20 possible points; 20% of responses were coded for interrater reliability, which was 87.5%. We report McDonald's omega as it tends to be more robust than Chronbach's alpha when scales might have responses that cluster around one end. Internal reliability was $\omega = 0.64$ in our sample.

Strange stories. The Strange Stories task presents brief vignettes that describe a social or physical event that participants are asked to explain (White et al., 2009). We focused on the eight social stories because they are designed to test the ToM construct. These stories describe interactions that feature a non-literal verbal utterance (e.g., a joke or an ironic statement). Participants read the scenarios in a binder followed by the test question by the experimenter, often asking why a character would say something (e.g., white lie). Responses were recorded and transcribed verbatim. Scoring followed White et al. (2009) and ranged from 0 to 2 for each story for a total of 16 possible points; 20% of responses were coded for interrater reliability, which was 92%. Internal reliability was $\omega = 0.48$ in our sample.

Higher-order false belief: In *An Imposing Memory Task* (Launay et al., 2015) participants read three short stories and then answered a series of true/false questions containing both factual recall and mentalizing questions, which vary from two to seven levels of embeddedness (the reader is considered level one). The task avoided ceiling effects because it included recursive reasoning beyond the second-order. For instance, one of the stories described a situation that required third-order reasoning. After reading the story, participants were asked whether it was true or false that one character thought that another character (Sam) was wrong in thinking a third character wanted to trick Sam. We were interested only in the mentalizing questions (10 per story for a total of 40 questions). Participants read and answered the questions in a Qualtrics survey. The Warnell and Redcay (2019) measurement, in comparison, contained 45 second and third-order ToM questions (adapted from Kinderman et al., 1998). We calculated Kuder–Richardson 20 (KR20) for internal reliability for the false belief task given its binary scoring (0 or 1), which was 0.69.

Four-item mentalizing index: This index was developed to assess self-reported mentalizing (i.e., non-emotional mental states; Clutterbuck et al., 2021). It included four statements, for example, *I find it easy to put myself in someone else's shoes*. Participants responded on a 4-point scale ranging from strongly disagree (1) to strongly agree (4) for a total of 16 possible points. Internal reliability for the four item mentalizing index is good; Clutterbuck et al. (2021) reported $\omega = 0.75$ and $\omega = 0.70$ in our sample. Given its questionnaire format, this task was also presented in the Qualtrics survey.

2.2.2 Covariates

Picture vocabulary test: Participants completed the Picture Vocabulary Test (PVT) using the NIH Toolbox application on an iPad. This task requires participants to select the picture (out of four) that best matches the word spoken. The PVT is an adaptive test; items differ across participants depending on their performance. Participants see a maximum of 25 items and administration stops when participants reach a standard error of less than 0.03. This measure has good test-rest reliability ($ICC = 0.80$) and converges with another measure of receptive vocabulary ($r = 0.80$) in a sample of adults up to the age of 85 (Gershon et al., 2014). Given the restricted age range of students in our study, age was unrelated to the PVT and uncorrected standard scores were used in analyses.

Executive function: Participants completed three executive function tasks from the NIH Toolbox application. In the Flanker Test, a measure of inhibitory control and attention, participants must focus on a directional stimulus in the center of the screen and respond with the correct direction (left/right) while inhibiting the stimuli flanking it. In the Dimensional Change Card Sort Test, a measure of cognitive flexibility and attention, participants must sort by both color and shape and correctly shift their attention to the current rule (color/shape). Finally, in the working memory task, participants must remember and sequence objects (animals or food) presented visually and via audio. In the 1-list version, participants see and hear a random sequence of either animals or food and must repeat them back in the order of their size. In the 2-list version, participants must repeat back both animals and food, sorting first by category then by size. Difficulty increases by 1 item (with two trials for each number of items) until participants incorrectly sequence both trials.

Scores in the Flanker and Dimensional Change Card Sort Test reflect both accuracy and reaction time given that adults tend to be highly accurate but slow down their reaction time to improve their accuracy. A log base 10 transformation is used to normalize the distribution of scores given that they tend to be positively skewed. This two-vector scoring allows interpretation of adult scores on these tasks in relation to performance of younger children (see Zelazo et al., 2014, for details). Scores in the working memory task reflect the number of correct items, transformed to a standardized score (Tulsky et al., 2014). The NIH Toolbox Flanker Test and Dimensional Change Card Sort Test (Zelazo et al., 2014) and working memory test (Tulsky et al., 2014) have been validated in adults (aged 20–85 years); they demonstrate high test–retest reliability and concurrent validity with other measures of executive function. Performance on these tasks tends to peak in early adulthood and decline in the later decades of life (Tulsky et al., 2014; Zelazo et al., 2014). Thus, these tasks are appropriate in adult samples. Given the restricted age range in the current sample, participant age was unrelated to any of the executive function measures and uncorrected standard scores were used in analyses.

Socioeconomic status: Because of the homogeneity of college students' education level, participants reported their parents' highest level of education on an 8-point scale ranging from less than high school degree (1) to professional degree (8); education level was averaged for students raised by two parents. Participants also reported their subjective social standing using the MacArthur Scale of Subjective Social Status which presents a ladder with 10 rungs in which 10 represents those with the most money, most education, and best jobs (Adler et al., 1994). This item was also presented in the Qualtrics survey. Z-scores of parents' average education level and subjective social status were summed to create an SES variable.

2.3 Procedure

Participants completed all assessments during two sessions in a university laboratory space one-on-one with a trained researcher. Informed consent was obtained during the first session followed by the Qualtrics survey (demographics, SES, the Higher-Order False Belief and FIMI tasks). The covariates were then administered on an iPad. The remaining ToM tasks (RMET, Strange Stories, and Animation) were completed at the second session.

2.4 Missing data

Eleven participants (5%) did not return for the second day of testing and so are missing day 2 measures (RMET, Strange Stories, and Animation). A few other variables are missing one to two participants due to experimenter error or noncompletion by the participant; thus, sample sizes range from 195 to 207 depending on the variable. All analyses were repeated with imputed data sets to utilize the full sample; all results were the same. Thus, results are presented with collected data only, not imputed data.

3 Results

Table 2 provides the descriptive statistics for all covariates and ToM measures. There was a wide distribution of parents' education; 20% of parents had a high school diploma or less; 35% had some college or an associate's degree; 28% had a bachelor's degree; and 17% had a graduate degree. The median (4.00) corresponds to an associate's degree. The mean for subjective social status was 6.13 (out of 10) and responses ranged from 1 to 9. There were no significant gender differences on any ToM measure ($ps > 0.33$) and so gender was not controlled in subsequent analyses. Despite the adult age group, there was wide variability on all ToM measures, demonstrating that the ToM skills tested were not at ceiling.

Next, we examined the bivariate correlations among all covariates and ToM measures (Table 3). Correlation coefficients with the covariates partialled out are in parentheses. Age and SES (i.e., the composite of parents' education and subjective social status) were unrelated to participants' ToM abilities. However, receptive vocabulary (PVT) was significantly and positively related to RMET, Strange Stories, and Higher-Order False Belief Understanding. Both of the ToM perceptual measurements (RMET and Animation) were related to a single EF component; otherwise, there were no relations between EF and ToM.

We next examined the factor structure of the five ToM tasks. We conducted an exploratory factor analysis, suppressing coefficients <0.40. Bartlett's test of sphericity was significant (35.39, *df* 10,

TABLE 2 Descriptive statistics for study variables.

Variable	*N*	Mean	SD	Range
Covariates				
Parent education	206	3.93	1.46	1–8
Subjective social status	207	6.13	1.58	1–9
Picture vocabulary test	205	99.16	6.96	76–118
Flanker	205	103.83	9.38	81–198
Dimensional change card sort	205	111.12	6.72	86–120
Working memory	205	106.34	9.85	78–136
Theory of mind				
Reading the mind in the eyes	196	7.32	1.59	3–10
Animation	196	12.88	3.73	2–20
Strange stories	195	12.12	2.27	5–16
Higher-order false belief	206	26.07	3.07	15–30
Four-item mentalizing index	207	12.83	2.09	6–16

TABLE 3 Bivariate and partial correlations among covariates and ToM.

	Age	SES	PVT	Flanker	DCCS	WM	RMET	AN	SS	FB	FIMI
Age	–	−0.06	0.13	−0.02	0.09	−0.05	0.00	0.02	−0.04	0.06	−0.08
SES		–	0.09	0.15*	0.09	0.08	−0.04	0.01	0.11	−0.10	0.08
PVT			–	0.08	0.18*	0.29***	0.25**	0.10	0.22**	0.25***	−0.12
Flanker				–	0.49***	0.07	0.15*	0.11	0.00	0.08	0.00
DCCS					–	0.13	0.04	0.17*	0.08	0.13	0.02
WM						–	0.07	0.12	0.09	0.08	−0.22**
RMET							–	0.07 (0.06)	0.08 (0.06)	0.11 (0.05)	−0.03 (0.01)
AN								–	0.18* (0.15*)	0.19** (0.17*)	0.16* (0.19*)
SS									–	0.27*** (0.26***)	0.10 (0.12)
FB										–	0.10 (0.14*)
FIMI											–

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. SES, socioeconomic status; PVT, Picture Vocabulary Test; DCCS, Dimensional change card sort; WM, Working memory; RMET, Reading the mind in the eyes test; AN, Animation; SS, Strange stories; FB, False belief; FIMI, Four-item mentalizing index. Correlations in parentheses control for age, SES, PVT, Flanker, DCCS, and WM.

$p < 0.001$). The Kaiser-Meyer-Olkin measure of sampling accuracy was 0.62. Two factors were identified based on eigenvalues (>1), total variance explained (52%), and examination of the scree plot. The eigenvalues for the first two factors were 1.54 and 1.03. The component matrix is presented in Table 4, showing that Animation, Strange Stories, and Higher-Order False Belief formed one factor, whereas Reading the Mind in the Eyes and the Four-Item Mentalizing Index formed a second factor but one in which the FIMI was negative.

Finally, we examined Factor 1 in relation to the covariates. A composite of the z-scores for Animation, Strange Stories, and Higher-Order False Belief was significantly related to receptive vocabulary ($r = 0.28$, $p < 0.001$), cognitive flexibility ($r = 0.19$, $p < 0.01$), and working memory ($r = 0.14$, $p < 0.05$), and unrelated to age, SES, or inhibitory control (all $ps > 0.23$).

TABLE 4 Factor loadings of ToM measures.

ToM measure	Factor 1	Factor 2
Reading the mind in the eyes		0.73
Animation	0.62	
Strange stories	0.67	
Higher-order false belief	0.69	
Four-item mentalizing index		−0.67

4 Discussion

There is a need in the field for (a) mapping out the subcomponents that comprise the ToM construct and (b) determining if broadly used

ToM tasks are measuring particular subcomponents of the ToM construct. Conceptual analysis is one approach toward addressing that need. Quesque and Rossetti (2020) identified a core component of the ToM construct—cognitive perspective taking—and then conceptually distinguished between widely used measurements that do or do not assess that core dimension. A second, complementary, approach in mapping out subcomponents is empirical in nature, and that is the one taken in this study.

This study investigated whether there is task coherence in early adulthood that maps onto the core cognitive perspective taking dimension posited by Quesque and Rossetti (2020). Couched within that question, we investigated the influence of individual differences in vocabulary, EF skills, and SES (which is a proxy for general cognitive abilities) on advanced ToM measurements. The results of the study carry four implications for our understanding of the ToM construct.

First, for the most part the correlations between different ToM measurements were very weak. That finding indicates ToM is not a unitary construct; in other words, there is not a single dimension of ToM that is captured across a spectrum of ToM instruments. Given the breadth of the construct, we suggest researchers avoid relying on a single measurement for advanced ToM and instead use a battery of instruments. Related, we also suggest that investigators consider examining different advanced ToM measurements separately in their analyses rather than forming a single composite that would obscure meaningful differences in what the instruments are measuring.

There is a caveat when interpreting the correlations. The internal consistency of the ToM measures was not strong, overall, in this study. That result is consistent with findings in samples of children and adolescents (e.g., Hayward and Homer, 2017; Osterhaus et al., 2016). The weak correlations could reflect, to some extent at least, the psychometric limitations of the instruments themselves. All else equal, low reliability reduces the likelihood of finding relations between instruments.

Second, with respect to cognitive influences, our results complement those found in the childhood research literature in that we found a stronger relationship between language and ToM than for EF and ToM (see Osterhaus and Bosacki, 2022). For the most part, measurements of three EF components were unrelated to ToM (see Table 3). It could be the case that EF skills have reached a level of maturity by early adulthood sufficient for mastering ToM task demands. Alternatively, it could be that ToM tasks with higher processing demands, such as tasks that use response time as a dependent variable, are more sensitive to EF differences than the untimed tasks administered in this study (see Apperly, 2021 for a discussion). A third possibility is that a different approach to measuring EF could result in a different finding. For instance, we assessed three different components of EF, but it could be that a battery of tasks focused on a single component (e.g., inhibition) yields a more robust link between ToM and adult EF than the one reported in this study. In sum, we have contributed evidence that speaks to the relation between EF and ToM in adulthood, but more research is needed to understand the magnitude and qualities of that relationship.

In contrast to EF, there is a clearer case for the relevance of language for ToM. As with most studies on advanced ToM, we used a measurement of receptive vocabulary to assess language skills (Osterhaus and Bosacki, 2022). We found weak, but non-negligible, relations between vocabulary skills and performance on tasks that place demands on verbal comprehension skills (see Table 3). We briefly elaborate on those relationships.

The Strange Stories and Higher-Order False Belief tasks each draw upon text comprehension skills for processing narratives in the tasks. Text comprehension is related to vocabulary knowledge (e.g., Ouellette, 2006). The association between the RMET and language (0.19 in magnitude) is worth noting because the RMET is often classified as a perceptual task. However, task performance also requires vocabulary knowledge of the various emotion and cognitive terms in the task (e.g., *pensive*). The inclusion of a glossary of definitions for the response options in the task reflects its relatively challenging vocabulary demands (see Miller, 2022). Considered as a whole, the results suggest that the relation between language and ToM observed in early childhood (see Milligan et al., 2007 for a meta-analysis) may be present in early adulthood, perhaps in a weakened form, on tasks that place a demand on verbal comprehension (vocabulary and reading). Returning to a question raised in the Introduction, the results suggest that even after basic language skills have reached maturity in early adulthood, individual differences may still relate to performance. As discussed earlier, the *expression* of advanced ToM knowledge on verbal tasks may vary in relation to vocabulary skills.

Third, if ToM is not a unified construct that demonstrates coherence among diverse measurements, is there evidence for a dimension, or subcomponent, of the construct? We found evidence that suggests different measurements overlap in assessing a dimension of the construct. The relation between Strange Stories and Higher-Order False Belief exceeded the threshold for a practically meaningful correlation in this study. The association was significant even after accounting for SES, language, and EF. The tasks are the two most widely used measurements of advanced ToM (Osterhaus and Bosacki, 2022), and their association is telling because they are highly representative of how the advanced ToM construct is operationalized in research. Recent effect size guidelines, including an analysis of over 700 meta-analytically derived correlations, characterize correlations between 0.2 and 0.3 as "typical" and "medium" effect sizes (Funder and Ozer, 2019; Gignac and Szodorai, 2016). We do not want to overstate the relation between Strange Stories and Higher-Order False Belief, but we propose that the relationship between them is theoretically meaningful.

In particular, the connection between Strange Stories and Higher-Order False Belief is consistent with Quesque and Rossetti (2020, p. 386) who argued that both instruments capture a core dimension of ToM: The "ability to corepresent—or to switch between—different perspectives." In essence, we found that (a) two instruments highly representative of the advanced ToM construct and theorized to tap into the perspective taking dimension of ToM are, in fact, related and (b) the relation is not due to shared covariance with other cognitive influences.

Fourth, unlike factor analytic results from Warnell and Redcay (2019), our results indicated an underlying ToM structure for advanced ToM. Two factors emerged in this study (see Table 4), but the second factor (RMET and FIMI) is uninterpretable. The RMET and FIMI were unrelated (see Table 3) and do not bear a clear conceptual connection because the FIMI was developed to measure the "understanding of non-emotional mental states" (Clutterbuck et al., 2021, p. 2). Although the FIMI and RMET were related in the Clutterbuck et al. study, Murphy et al. (2022) point out that if the RMET is primarily measuring empathy and emotion perception, then it should actually be unrelated to the FIMI. The findings in the present study are consistent with that interpretation. The age range (18–78-years) in Clutterbuck et al. (2021) was substantially different from the more restricted range in this study. ToM proficiency changes throughout the lifespan (e.g., Miller, 2022), and it could be that there is a developmental explanation for the different results.

Factor 1, however, is interpretable based upon previous evidence and theory. The Strange Stories, Higher-Order False Belief, and Animation tasks formed that factor. The inclusion of the Strange Stories and False Belief measures within a single factor supports claims that both instruments measure perspective taking about cognitive states (Quesque and Rossetti, 2020). The inclusion of the Animation task is inconsistent with Quesque and Rossetti who argued the task does not entail representing mental states. They proposed that performance on the task is better explained by low-level processes such as kinematic processing and visual discrimination. Those processes trigger the perception of intention, which is dissociable from the representation of, and reasoning about, covert mental states (Gobbini et al., 2007).

Based upon a hierarchical clustering of meta-analytic results from neuroimaging data in adults, Schurz et al. (2021) proposed that animation tasks represent an intermediate dimension, one that overlaps with affective and cognitive dimensions of social cognition. The association with affective tasks (e.g., measurements of empathy) occurs because the geometric animations depict scenarios that trigger emotional reactions. The overlap with the cognitive dimension occurs because animation tasks also involve inference and reasoning about the mental states of agents. The inclusion of the Animation task in Factor 1 is more closely aligned with the Schurz et al. (2021) analysis of the animation task than with Quesque and Rossetti (2020).

The different loadings of the RMET and Animation tasks are interesting because the tasks appear to share two features in common. One, as just noted, the Animation task presents scenarios that lend themselves to emotionally laden interpretations. In that regard, both tasks are apparently tapping into the affective dimension of ToM. Second, the attribution of mental states and emotions for both tasks have a perceptual basis. In that sense, both tasks appear to tap into social-perceptual processes that infer mental states from nonverbal cues (see Meinhardt-Injac et al., 2020 for a discussion of a social-perceptual ToM component). That is, emotion, thoughts, and intentions are attributed based upon what one observes from geometric movement or the eye regions rather than upon verbal information. Despite their similarities, in the next section we suggest the tasks load onto different factors because only the Animation task has a narrative structure.

Before discussing Factor 1, it is worth highlighting that there is a preliminary, uncertain, quality to it given the small correlations between the instruments that comprise it. In our discussion in the next section, we treat the factor as grounds for future research rather than as conclusive evidence for an underlying structure of ToM.

4.1 Narrative comprehension as a cognitive influence on ToM?

Quesque and Rossetti (2020, p. 386) characterized the ability to represent different perspectives as a "core component" of ToM. However, the nature of Factor 1 suggests that their characterization is incomplete in failing to account for the *purpose* for perspective taking. Participants identify and reason about perspectives in order to make sense of what others say and do. Ho et al. (2022, p. 959) state that, "The classic problems (and psychological tasks) used to study ToM require an observer to predict or explain another person's action." In all three tasks in Factor 1, participants are interpreting and explaining the actions of agents by inferring mental states within a narrative structure. In that respect, the tasks differ from both the FIMI and RMET. Also in that respect, the three tasks in Factor 1 have in common the requirement that participants reason about mental states within the causal-explanatory framework discussed in the Introduction as a central dimension of ToM.

We propose that the narrative framework of the tasks that comprise Factor 1 helps explain their overlap. The narrative structure of the Strange Stories and Higher-Order False Belief tasks is clear because each task features written text that portrays story-like social scenarios. In contrast, the story structure is inferred by participants from perceptual stimuli in the Animation task. Devine et al. (2016) argued that Strange Stories and Animation tasks both involve inferring mental states to interpret social scenarios. In that characterization (which we share), the Animation task is not simply the low-level recognition of mental states from perceptual stimuli. Rather than passively observing the movement of geometric shapes, participants' active interpretations appear to invoke a story schema; that is, a memory for how stories are structured into common narrative elements of problem, goal, action, and resolution (e.g., Meadowcroft and Reeves, 1989; see also, Jing and Kirkorian, 2020). Consistent with that interpretation, Nguyen et al. (2019) found that neural activity in regions of the brain associated with narrative and linguistic processing was elicited when adult participants watched nonverbal, story-like animated scenarios of geometric shapes.

Narrative comprehension involves the integration of events within a story while drawing upon one's own prior experiences and knowledge to interpret them. Participants presumably draw upon their own social knowledge and experiences when explaining what happened in the animated scenarios. For instance, nearly 20% of the participants in this study described a scene in the Animation task (when an agent "surprises" another) by referring to a social ritual—a prank that involves knocking on a door and running away ("ding dong ditch"). Thus, despite the difference in modality—written text or animated motion—all three tasks require integrating a series of events into a coherent structure, while drawing upon one's own experiences and ToM knowledge to interpret behavior in the scenarios. Only with this prior narrative structure would one extrapolate a social ritual from the movements of two triangles and an empty box.

We refer to Factor 1 as the *protagonist perspective* factor following from Mason and Just (2009). They characterized the overlapping neural regions that support ToM reasoning and narrative comprehension as the *protagonist perspective neural network*. The term refers to the processes used to interpret the action of "any human, animal, or other entity capable of autonomous action that is the focus of a story" (p. 158). The neural network is associated with two neurocognitive processes. One involves monitoring the narrative itself (e.g., monitoring what the protagonist says or does in the context of successive events). The second involves integrating the narrative events in order to make a mentalistic inference to predict and explain action. Both processes are relevant to the tasks in the protagonist perspective factor and neither one is relevant to the two tasks (RMET and FIMI) not included in the factor. In short, the tests in the factor involve reasoning about "perspective" *and* "protagonists" (i.e., the actions of a protagonist in the context of a story-like scenario).

The protagonist perspective factor was positively associated with vocabulary knowledge. The magnitude of the correlation, 0.28, exceeded our threshold of a practically meaningful relationship (below that threshold, there were also very weak correlations with EF measurements too). Vocabulary knowledge is important for accessing background world knowledge that supports narrative inferences and comprehension (e.g., Cromley and Azevedo, 2007). It could be the case that narrative comprehension skills, supported by vocabulary, are a cognitive influence on the tasks in the protagonist perspective factor. Two EF components (working memory and cognitive flexibility) each had a weak but statistically significant relationship with the composite of the z-scores for Animation, Strange Stories, and Higher-Order False Belief. In the Introduction, we discussed that individual differences in general cognitive variables could lead to variability in the expression of ToM knowledge in early adulthood. EF processes such as keeping the details of narrative elements in mind might also exert a weak, but non-negligible, influence on the expression of ToM knowledge on narrative-based tasks in early adulthood (see Cartwright et al., 2020 for a discussion of EF and reading comprehension in early adulthood).

4.2 Limitations and future research

There is considerable variation in the response demands of advanced ToM measures, which attenuates correlations and complicates attempts to identify if they conceptually cohere (Devine, 2021). There is a need for research that isolates the influence of task demands, doing so by (a) administering a battery of tasks that share similar conceptual demands, while (b) systematically varying their response demands. Doing so would help disentangle the influence of task-related and conceptual variables on task associations. Limitations in internal consistency among ToM measurements is also a point of caution, and those limitations highlight a need for replication.

Often the internal consistency of ToM measurements is not reported in the literature (Osterhaus and Bosacki, 2022), which complicates basing task selection upon psychometric properties. When possible, we attempted to select improved measures (e.g., the abbreviated RMET, Olderbak et al., 2015). In this study, our goal was to select measures because of their wide usage and their theoretical relevance to the claim that cognitive perspective taking is a core ToM dimension. The instruments are widely used because they are related to conceptually-relevant social and academic variables, an outcome that is evidence of their predictive validity (see Miller, 2022 for a comprehensive review). They are highly representative of the advanced ToM construct.

Researchers routinely use all of the items on a ToM test or use a shortened version based upon a consideration other than reliability (e.g., subtests based upon item content). Practically speaking, researchers usually select a task, and not task items, in ToM research. Following from that practice, our focus was upon taking existing instruments in the field and examining whether they are related, in conceptually meaningful ways, to other ToM instruments and measures of general cognitive variables. Our focus was not on testing the psychometric properties of each one. Nevertheless, we acknowledge that because of relatively weak internal consistency, a ToM instrument could be tapping more than one ToM dimension. A promising alternative approach to the one adopted in this study would involve conducting a factor analysis on individual items (e.g., Osterhaus et al., 2016). In doing so, future research on advanced ToM in adulthood might uncover factors that were not revealed by the approach used in this study.

As we noted in the Introduction, the ToM instruments were chosen for this study because they are widely used, and because they are found in research with both children and adults. While those considerations are consistent with the rationale for the study, the result is that newer, promising measurements were not included. Of particular interest are instruments that have recently refined measurement of the construct in ways that are evident at the neural level. In research with adults, measurements such as the EmpaToM (Kanske et al., 2015) isolate the constructs of ToM and empathy, and find distinct neural profiles associated with each construct (see also, Völlm et al., 2006). ToM instruments associated with specific neural correlates are particularly well-situated to advance a research agenda that aligns the construct, its measurement, and underlying cognitive and neural processes.

Future research is also needed to determine if the protagonist perspective factor represents an underlying structural dimension of ToM. Our sample size was comparable to Warnell and Redcay (2019), but research with a larger sample size would increase confidence in the replicability of the factor. Adding new instruments to the ToM battery – which would add to the need for a larger sample size—is also an area for future research. For instance, the significant relation between the factor and vocabulary clearly signals a role for language skills. Consequently, additional tasks that significantly depart from the verbal demands found in the protagonist perspective tasks (e.g., comprehending verbal narratives, and articulating explanations for characters' actions) would likely fall outside of the factor. Related, a determinative variable for whether a task "belongs" in a protagonist perspective factor is the extent to which the task requires narrative processing over a series of events. Variability on that dimension should be linked to a task's inclusion or exclusion in a protagonist perspective factor.

More generally, there is also a need for research on the role that narrative processing plays in ToM task performance. Narrative comprehension involves general inferencing abilities and the integration of information into a coherent, causally-related sequence of events. More research is needed to pinpoint how (or whether) those skills also support ToM reasoning, particularly in light of recent meta-analytic evidence that finds a relation between text comprehension and ToM across the lifespan (Tompkins et al., 2024). Evidence that preschoolers' ToM is related to their causal sequencing of narratives is a promising step in that regard (Tompkins et al., 2020).

We only included one measurement in this study, the RMET, identified by researchers as representative of an affective dimension of ToM (e.g., Schurz et al., 2021). Therefore, our findings do not speak to that hypothesized feature of the ToM construct. It is worth asking in future research if narrative-like scenarios designed to elicit perspective taking about affective states might align more closely with the RMET or with tasks associated with the protagonist perspective factor. In other words, is the mental state content that participants attribute (affective or cognitive) more determinative than the purpose of those attributions (to predict and explain action in a narrative structure) when assessing task coherence?

Our study was also limited to a restricted age range. In doing so, we avoided conflating individual and developmental differences. But the results do not speak to whether there is continuity or discontinuity in ToM development. Recall from the Introduction that we intentionally chose instruments in this study that are also used in child development studies. In doing so, our findings pinpoint a question for future research. Osterhaus et al. (2016) reported a "social reasoning" factor in children ages 8- to 10-years that was comprised of Higher-Order False Belief measurements, a subset of items from the Strange Stories task, and a child version of the RMET. Whether the different findings with respect to the RMET reflect developmental discontinuity or methodological differences between studies requires further research. As we discussed in the Introduction, it could be the case that ToM in childhood is less differentiated than it is in adulthood; consequently, distinctions between measurements such as Higher-Order False Belief and RMET may be found in adults but not in childhood. In sum, this study provides evidence for how the instruments align in adulthood. Evidence for continuity and discontinuity in the structure of ToM awaits research with different age groups.

Alternatively, an expanded focus that includes older participants would also inform how (or whether) response demands on cognitive abilities influence task coherence. For instance, there is evidence that age-related declines in ToM are tied to declines in EF rather than ToM reasoning (Cho and Cohen, 2019). When task-related EF demands are reduced, age-related deficits in ToM performance are also reduced. The point, more generally, is that the likelihood of coherence could be tied to moderators such as age that influence whether demands mask conceptual knowledge. Similar concerns extend to neurodivergent individuals for whom performance factors may substantially obscure underlying conceptual competencies (see Montgomery et al., 2023 for a discussion). The extent to which the findings reported in this study are generalizable to other populations awaits further research.

5 Conclusion

In sum, this research investigated the nature of the ToM construct by asking if commonly used advanced ToM measures cohere in early adulthood. We also addressed a need in the literature for a comprehensive analysis of the influence of vocabulary, EF, and SES on adult ToM. We administered a battery of tasks to over 200 participants who, collectively, mirrored race and ethnicity proportions in the U.S. The Strange Stories and Higher-Order False Belief task exhibited a meaningful association. We also found evidence for what we termed a protagonist perspective factor. The factor is related to general vocabulary skills, indicating that the influence of language on ToM extends beyond childhood. The factor also implicates narrative comprehension skills as an additional cognitive influence on ToM, and that suggestion warrants future research.

Data availability statement

The raw data supporting the conclusions of this article will be made available upon reasonable request.

Ethics statement

The studies involving humans were approved by the Ohio State University Institutional Review Board. 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

DM: Conceptualization, Investigation, Methodology, Writing – original draft. VT: Conceptualization, Formal analysis, Investigation, Methodology, Writing – review & editing. XF: Formal analysis, Methodology, Writing – review & editing.

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.

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

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., et al. (1994). Socioeconomic status and health: the challenge of the gradient. *Am. Psychol.* 49, 15–24. doi: 10.1037/0003-066X.49.1.15

Apperly, I. (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 (London: Routledge Press), 37–54.

Apperly, I. A., Samson, D., and Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Dev. Psychol.* 45, 190–201. doi: 10.1037/a0014098

Apperly, I. A., and Wang, J. J. (2021). "Mindreading in adults: cognitive basis, motivation, and individual differences" in The cognitive basis of social interaction across the lifespan. eds. H. J. Ferguson and E. E. F. Bradford (Oxford: Oxford University Press), 96–116.

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

Carlson, S. M., Koenig, M. A., and Harms, M. B. (2013). Theory of mind. *WIREs Cognit. Sci* 4, 391–402. doi: 10.1002/wcs.1232

Cartwright, K. B., Lee, S. A., Taboada Barber, A., DeWyngaert, L. U., Lane, A. B., and Singleton, T. (2020). Contributions of executive function and cognitive intrinsic motivation to university students' reading comprehension. *Read. Res. Q.* 55, 345–369. doi: 10.1002/rrq.273

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:e0222890. doi: 10.1371/journal.pone.0222890

Clutterbuck, R. A., Callan, M. J., Taylor, E. C., Livingston, L. A., and Shah, P. (2021). Development and validation of the four-item mentalising index. *Psychol. Assessm.* 33, 629–636. doi: 10.1037/pas0001004

Cromley, J. G., and Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *J. Educ. Psychol.* 99, 311–325. doi: 10.1037/0022-0663.99.2.311

de Villiers, J. (2021). With language in mind. *Lang. Learn. Dev.* 17, 71–95. doi: 10.1080/15475441.2020.1820338

Devine, R. T. (2021). "Individual differences in theory of mind in middle childhood and adolescence" in Theory of mind in middle childhood and adolescence. eds. S. Lecce and T. D. Devine (London: Routledge Press), 55–76.

Devine, R. T., and Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: a meta-analysis. *Child Dev.* 85, 1777–1794. doi: 10.1111/cdev.12237

Devine, R. T., and Hughes, C. (2018). Family correlates of false belief understanding in early childhood: a meta-analysis. *Child Dev.* 89, 971–987. doi: 10.1111/cdev.12682

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, 758–771. doi: 10.1037/dev0000105

Funder, D. C., and Ozer, D. J. (2019). Evaluating effect size in psychological research: sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* 2, 156–168. doi: 10.1177/2515245919847202

Gershon, R. C., Cook, K. F., Mungas, D., Manly, J. J., Slotkin, J., Beaumont, J. L., et al. (2014). Language measures of the NIH toolbox cognition battery. *J. Int. Neuropsychol. Soc.* 20, 642–651. doi: 10.1017/S1355617714000411

Gignac, G. E., and Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personal. Individ. Differ.* 102, 74–78. doi: 10.1016/j.paid.2016.06.069

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., and Haxby, J. V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *J. Cogn. Neurosci.* 19, 1803–1814. doi: 10.1162/jocn.2007.19.11.1803

Happé, F., Cook, J. L., and Bird, G. (2017). The structure of social cognition: in (ter) dependence of sociocognitive processes. *Annu. Rev. Psychol.* 68, 243–267. doi: 10.1146/annurev-psych-010416-044046

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

Ho, M. K., Saxe, R., and Cushman, F. (2022). Planning with theory of mind. *Trends Cogn. Sci.* 26, 959–971. doi: 10.1016/j.tics.2022.08.003

Jing, M., and Kirkorian, H. L. (2020). Teaching with televised stories: a story-focused narrative preview supports learning in young children. *Child Dev.* 91, e1101–e1118. doi: 10.1111/cdev.13385

Kanske, P., Böckler, A., Trautwein, F. M., and Singer, T. (2015). Dissecting the social brain: introducing the EmpaToM to reveal distinct neural networks and brain–behavior relations for empathy and theory of mind. *NeuroImage* 122, 6–19. doi: 10.1016/j.neuroimage.2015.07.082

Karr, J. E., Rodriguez, J. E., Goh, P. K., Martel, M. M., and Rast, P. (2022). The unity and diversity of executive functions: a network approach to life span development. *Dev. Psychol.* 58, 751–767. doi: 10.1037/dev0001313

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

Klindt, D., Devaine, M., and Daunizeau, J. (2017). Does the way we read others' mind change over the lifespan? Insights from a massive web poll of cognitive skills from childhood to late adulthood. *Cortex J. Dev. Study Nerv. Syst. Behav.* 86, 205–215. doi: 10.1016/j.cortex.2016.09.009

Launay, J., Pearce, E., Wlodarski, R., van Duijn, M., Carney, J., and Dunbar, R. I. (2015). Higher-order mentalising and executive functioning. *Personal. Individ. Differ.* 86, 6–14. doi: 10.1016/j.paid.2015.05.021

Lecce, S., Bianco, F., and Hughes, C. (2021). Reading minds and reading texts: evidence for independent and specific associations. *Cogn. Dev.* 57:101010. doi: 10.1016/j.cogdev.2021.101010

Mason, R. A., and Just, M. A. (2009). The role of the theory-of-mind cortical network in the comprehension of narratives. *Lang. Ling. Compass* 3, 157–174. doi: 10.1111/j.1749-818X.2008.00122.x

Meadowcroft, J. M., and Reeves, B. (1989). Influence of story schema development on children's attention to television. *Commun. Res.* 16, 352–374. doi: 10.1177/009365089016003003

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. Oxford: 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

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

Montgomery, D. E., Tompkins, V., and Beningo, J. (2023). Theory of mind. ed. J. L. Matson. *Handbook of Clinical Child Psychology.* Springer.

Murphy, B. A., Hall, J. A., and Duong, F. (2022). It looks like construct validity, but look again: comment on Clutterbuck et al. (2021) and recommendations for test developers in the broad "empathy" domain. *Psychol. Assess.* 34, 397–404. doi: 10.1037/pas0001063

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

Nguyen, M., Vanderwal, T., and Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage* 184, 161–170. doi: 10.1016/j.neuroimage.2018.09.010

Olderbak, S., Wilhelm, O., Olaru, G., and Geiger, M. (2015). A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Front. Psychol.* 6:155797. doi: 10.3389/fpsyg.2015.01503

Osler, M., Avlund, K., and Mortensen, E. L. (2013). Socio-economic position early in life, cognitive development and cognitive change from young adulthood to middle age. *Eur. J. Pub. Health* 23, 974–980. doi: 10.1093/eurpub/cks140

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., Koerber, S., and Sodian, B. (2016). Scaling of advanced theory-of-mind tasks. *Child Dev.* 87, 1971–1991. doi: 10.1111/cdev.12566

Ouellette, G. P. (2006). What's meaning got to do with it: the role of vocabulary in word reading and reading comprehension. *J. Educ. Psychol.* 98, 554–566. doi: 10.1037/0022-0663.98.3.554

Peloquin, C. M., McDermott, C. H., and Moses, L. J. (2023). Individual differences in second-order false-belief understanding and executive abilities: a meta-analytic review of evidence from school-age children and adults. *Cogn. Dev.* 68:101367. doi: 10.1016/j.cogdev.2023.101367

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

Quesque, F., and Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspect. Psychol. Sci.* 15, 384–396. doi: 10.1177/1745691619896607

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nat. Rev. Psychol.* 1, 223–235. doi: 10.1038/s44159-022-00037-z

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., and Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends Cogn. Sci.* 19, 65–72. doi: 10.1016/j.tics.2014.11.007

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., et al. (2021). Toward a hierarchical model of social cognition: a neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychol. Bull.* 147, 293–327. doi: 10.1037/bul0000303

Tompkins, V., Blosser, M. K., and Downing, M. (2020). False belief understanding and narrative comprehension in the preschool years. *Cogn. Devel.* 56, 100936. doi: 10.1016/j.cogdev.2020.100936

Tompkins, V., Farrar, M. J., and Montgomery, D. E. (2019). Speaking your mind: language and narrative in young children's theory of mind development. *Child Dev. Behav.* 56, 109–140. doi: 10.1016/bs.acdb.2018.11.003

Tompkins, V., Logan, J. A., Blosser, D. F., and Duffy, K. (2017). Child language and parent discipline mediate the relation between family income and false belief understanding. *J. Exp. Child psychol.* 158, 1–18. doi: 10.1016/j.jecp.2017.01.001

Tompkins, V., Logan, J. A., Blosser, D. F., and Duffy, K. (2024). Theory of mind and text comprehension across the lifespan: A meta-analysis. *Develop. psychol.* doi: 10.1037/dev0001869 [Online Publications].

Tulsky, D. S., Carlozzi, N., Chiaravalloti, N. D., Beaumont, J. L., Kisala, P. A., Mungas, D., et al. (2014). NIH toolbox cognition battery (NIHTB-CB): list sorting test to measure working memory. *J. Int. Neuropsychol. Soc.* 20, 599–610. doi: 10.1017/S135561771400040X

U.S. Census Bureau (2022). Quick facts: United States. Available at: https://www.census.gov/quickfacts/fact/table/US/PST045222(Accessed April 5, 2024).

Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., et al. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *NeuroImage* 29, 90–98. doi: 10.1016/j.neuroimage.2005.07.022

Warnell, K. R., and Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition* 191:103997. doi: 10.1016/j.cognition.2019.06.009

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

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684. doi: 10.1111/1467-8624.00304

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

White, S., Hill, E., Happé, F., and Frith, U. (2009). Revisiting the strange stories: revealing mentalizing impairments in autism. *Child Dev.* 80, 1097–1117. doi: 10.1111/j.1467-8624.2009.01319.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

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., Conway, K. P., et al. (2014). NIH toolbox cognition battery (CB): validation of executive function measures in adults. *JINS* 20, 620–629. doi: 10.1017/S1355617714000472

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 19 December 2024
DOI 10.3389/fpsyg.2024.1453140

Check for updates

OPEN ACCESS

EDITED BY
Ann Dowker,
University of Oxford, United Kingdom

REVIEWED BY
Antonio Narzisi,
Stella Maris Foundation (IRCCS), Italy
Meng-Jung Liu,
National Kaohsiung Normal University, Taiwan

*CORRESPONDENCE
Fu-Qiang Qiao
✉ sep_qiaofq@ujn.edu.cn

RECEIVED 24 June 2024
ACCEPTED 25 November 2024
PUBLISHED 19 December 2024

CITATION
Qiao F-Q, Li S-N, Du T-T, Cheng W-M, Sun Y-Y, Qiang X, Dong Y-J and Wang L (2024) Spanning three decades: global research wave and future prospects of broader autism phenotype—a visual researches by CiteSpace and VOS viewer.
Front. Psychol. 15:1453140.
doi: 10.3389/fpsyg.2024.1453140

Spanning three decades: global research wave and future prospects of broader autism phenotype—a visual researches by CiteSpace and VOS viewer

Fu-Qiang Qiao[1]*, Si-Ning Li[1], Tong-Tong Du[1], Wen-Ming Cheng[2], Ying-Ying Sun[3], Xu Qiang[4], Ya-Jie Dong[1,5] and Lei Wang[6]*

[1]School of Education and Psychology, University of Jinan, Jinan, China, [2]Department of Wei Zhong Children's Rehabilitation Center, Jinan, China, [3]School of Philosophy and Social Development, Shandong University, Jinan, China, [4]Department of Psychology, Ningbo University, Ningbo, China, [5]Tianjin Yang Jialou Senior High School, Tianjin, China, [6]Department of Rehabilitation Medicine, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, China

Objectives: To conduct a comprehensive review of the literature pertaining to the broader autism phenotype, the paper endeavors to delineate the key research directions and topics, document the current research trends, and furnish insightful analyses and novel perspectives to foster future advancements in the field, with the aid of CiteSpace and VOS viewer.

Methods: CiteSpace and VOS viewer are two kinds of software for visualizing citations that is intended to examine academic literature and identify possible sources of knowledge. The Web of Science Core Collection database was used to retrieve articles from 1994 to 2024 that discussed the autism phenotype in general. Following data collection, analyses were conducted using CiteSpace V.6.2.R4 (64-bit) on a number of topics, such as annual publication output, highly cited journals, affiliations with countries and institutions, eminent authors, cited references, and keywords. Knowledge maps, collaborative network analysis, cluster analysis, and strongest citation burst analysis were among the tools used to visualize the data.

Results: A total of 1,075 articles about the broader autism phenotype were found. Roughly speaking, the annual number of publications is rising. With 546 articles on the subject, the United States is the nation with the greatest amount of authority and influence (centrality of 0.32), with England (218 articles) and Canada (115 articles) coming in second and third, respectively. The cited journals and institutions are mainly from the United States. The research consistently emphasizes the academic achievement and engagement of broader autism phenotype when analyzing the most cited references and authors.

Conclusion: This study used CiteSpace to analyze the state of the larger autism phenotype field and then identified research frontiers and hotspots. As new findings are made, global trends in more thorough studies of the autism phenotype suggest that interest in these studies will only grow.

KEYWORDS

broader autism phenotype, bibliometric analysis, CiteSpace, research trends, web of science, VOS viewer, visual analysis

1 Introduction

The broader autism phenotype (BAP) is a subclinical set of personality and other features that is thought to index familiarity and/or genetic liability to autism (Lainhart et al., 2002). This view suggests that BAP is milder but qualitatively similar to the diagnostic autism phenotype. A tendency toward rigidity (lack of interest in and/or difficulty adjusting to change) and/or hypersensitivity (excessive distress at remarks or behavior of others that is felt to be critical or insensitive), abnormalities in speech, and qualitatively impaired friendships were considered criteria for the BAP (Lainhart et al., 2002). BAP traits were first observed by Kanner (1943), he implicated parents as a potential cause for autism, claiming that most parents of the children studied did not seem exceptionally warm to their children. Twin and family studies have confirmed that the development of autism is highly heritable (Bolton et al., 2010; Bailey et al., 1995). The twin study of autism by Folstein and Rutter (1977) was the first systematic study to suggest the existence of a broad autism phenotype. Many studies have been done on the impact of the parents' broader autism phenotype, which is a collection of subclinical traits of autism spectrum disorders, on the child's autistic disorder phenotype (Wheelwright et al., 2010). As of right now, no single paradigm can adequately explain the nature of BAP and its rising incidence. Several studies have confirmed that the broad phenotype is mainly characterized by cognitive deficits, social deficits, stereotypical behaviors, but also by anxiety, depression, hypersensitivity to criticism, poor adaptability, and over-responsiveness, with social deficits appearing to be more important (Baron-Cohen et al., 2001; Sung et al., 2005; Hurley et al., 2007; Kanne et al., 2012). In conclusion, the above studies confirm the existence of broad autism phenotype and that some relatives of autistic pre-diagnostics have subclinical symptoms that resemble autistic-like manifestations.

Bibliometrics is a quantitative approach to analyzing scholarly publications, which has been applied in many areas of research to assess the patterns of authors, institutions, journals, countries and keywords et al. associated with particular types of publication (Liang et al., 2017). It is also a good choice for identifying research trends and knowledge gaps of a research field over time (Wu et al., 2021). CiteSpace, as one of visualization analysis software, is widely used to explore the research hotspots, research frontiers, knowledge base, main authors and institutions of a research field, as well as to help predict the future direction of a research field (Chen et al., 2022). However, there is a lack of summary and evaluation of the characteristics of the literature, research directions, depth of research, and research hotspots in BAP research. Therefore, there is a need to determine the current state of BAP to inform future researches. In this study, it was aimed to analyze the global research trends and prospects on the BAP last 13 years, CiteSpace was applied to make a bibliometric analysis of related articles collected from Web of Science Core Collections database from 1994 to 2024. Patterns of research publications in this field were mapped to author, institutions, journals, countries, keywords, references, research themes, research hotspots and emerging research areas regarding BAP.

2 Methods

2.1 Data source

As the independent global citation database for the most reputable publishers in the world, Web of Science (WoS) is the premier research platform for information in the natural sciences, social sciences, arts, and humanities (Luo et al., 2021; Zhou et al., 2021). Additionally, papers that are part of the Web of Science Core Collection (WoSCC) are regarded as being crucial to the research process (Wu et al., 2021). For this study, we chose WoSCC database as our primary data source.

2.2 Search strategy

We collected data on WoSCC on October 31, 2024.The four components of the data retrieval strategy were as follows: The topics are "broader autism phenotype," "article" or "review article," and "language" is "English." Complete records were downloaded in plain text format for additional analysis, along with the cited references that corresponded with them. In Figure 1, the flowchart is displayed.

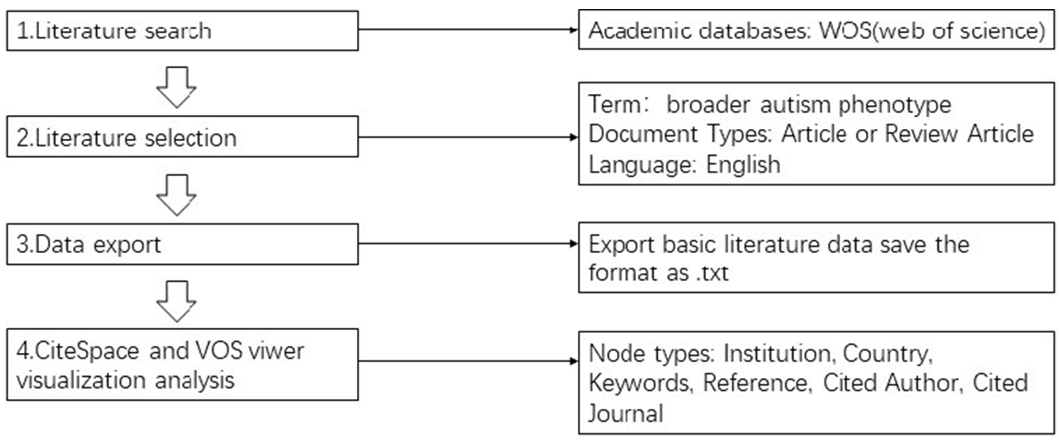


FIGURE 1
The flowchart of data selection.

2.2.1 Inclusion criteria

Inclusion criteria were: (i) articles on BAP, limited to English as a language; (ii) articles taken from the WoSCC database.

2.2.2 Exclusion criteria

The following were the exclusion criteria: (i) hand-and telephone-collected articles, or newsletters, notices, announcements, calls for papers, and conference papers; (ii) articles not written in English; (iii) conference proceedings and abstracts; (iv) duplicate publications or the same study; (v) unpublished articles and unrelated articles.

2.3 Analysis tool

CiteSpace, a bibliometric analysis software developed by Chen (2004), is a citation visual analysis tool that enables the exploration of knowledge potential in scientific literature and gradually evolves under the background of scientometrics and data visualization (Gao et al., 2022). This study utilized CiteSpace V.6.2.R4 (64-bit) to analyze relevant research on BAP, The objective was to provide evidence-based support for educators and researchers, gain insights into the current state and trends in the field, and generate new ideas for future development.

VOSviewer is a free and open-source software for visualizing and analyzing scientific literature, keywords and author relationships, developed by Leiden University in the Netherlands. It is a knowledge graph visualization tool mainly used to build and view bibliometric knowledge graphs. The software is based on the principle of co-citation and co-citation of literature, has strong visualization capabilities, is suitable for large-scale sample data, and supports a variety of view browsing methods, such as label view, density view, clustering view, and dispersion view.

R language was originally developed by two people from the University of Auckland, New Zealand, Ross Ihaka and Robert Gentleman, R language has a wide range of applications, this paper focuses on the application of R language for data analysis and visualization, statistical analysis, data mining, machine learning and other work, and through a variety of charts and visualization means to show the results of data analysis.

2.4 Data analysis

CiteSpace software was used in this study to find the citation bursts in a number of different dimensions, such as hotspots, keywords, research publication year, author, research institution, journal, and country. With the help of CiteSpace, one may create a visual knowledge network made up of nodes and connections. The nodes stand for various aspects including authors, countries, institutions, and cited references, while the linkages connecting the nodes show collaboration or co-citation relationships. Nodes' sizes indicate how frequently or how much they occur, and different colors stand for different years—lighter hues denoting more recent years, and darker colors representing earlier years. Purple circles also indicate centrality. High centrality nodes are frequently seen as pivotal or turning points in the field (Liu et al., 2019).

3 Results

3.1 Annual publication analysis

After applying the search strategy, 1,075 publications meeting the inclusion criteria were found; Figure 2 displays the annual number of articles published. It's evident that 2018, 2019 and 2023 are the years with the most publications. The number of publications has been trending in a fairly stable direction in recent years. The fact that the majority of the publications in this field were published within the previous ten or so years further indicates how new this field is.

3.2 Countries analysis

As shown in Figure 3A, these publications come from more than 70 countries/regions. The top ten countries/regions

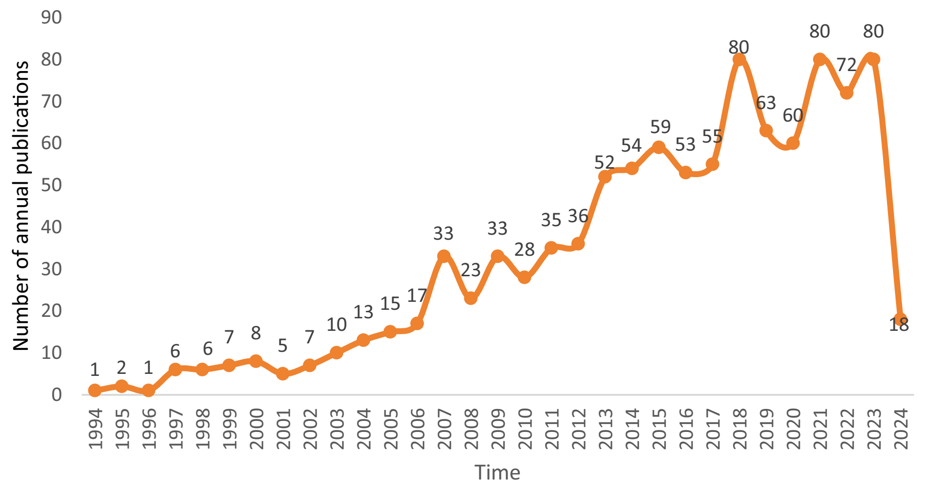


FIGURE 2
The number of publications from 1994 to March April 2024.

10.3389/fpsyg.2024.1453140

produced 99.16% (1,066/1075) of the publications, with USA (50.79%, 546/1075), the England (20.27%, 218/1075), and Canada (10.69%, 115/1075) leading the way. We used 1-year slices to analyze all BAP articles published between 1994 and 2024 in order to look into the relationships between articles published in various countries. A merged network with 70 nodes and 541 links, representing a country-specific distribution map, was produced by this analysis and is shown in Figure 3B. The differences in the number of publications of each country can also be seen relatively directly from the size of the circles in Figure 3C. Countries are represented by the nodes, and their relationships are shown by the lines. Larger nodes indicate higher publication productivity. The size of the nodes reflects the number of publications. Table 1 lists the twenty most productive nations in the world in terms of BAP.

As we can see, the United States is leading all other countries by a significant margin and has the highest publications productivity (546). It is also the only nation with over 200 articles. England, Canada, Australia, and Italy came next.

3.3 Cited journals analysis

Through an examination of the journals cited in the field of BAP, we can acquire valuable insights into the primary sources of knowledge dissemination and effectively identify pertinent data. We used 1-year slices to analyze all BAP articles published between 2011 and 2024, classifying the cited journals as nodes. This resulted in a distribution map of cited journals with a merged network that had 1,260 nodes and 5,392 links, as seen in Figure 4. Journals are represented by nodes, and the relationships between journals are shown by lines. Higher frequency nodes are typically regarded as significant nodes that have a bigger influence on how a scientific field develops. A tighter relationship is indicated by a thicker line connecting two nodes.

3.4 Institutions analysis

In order to examine cooperative relationships between institutions and pinpoint significant institutions in the field of BAP, we created a network map using Cite.

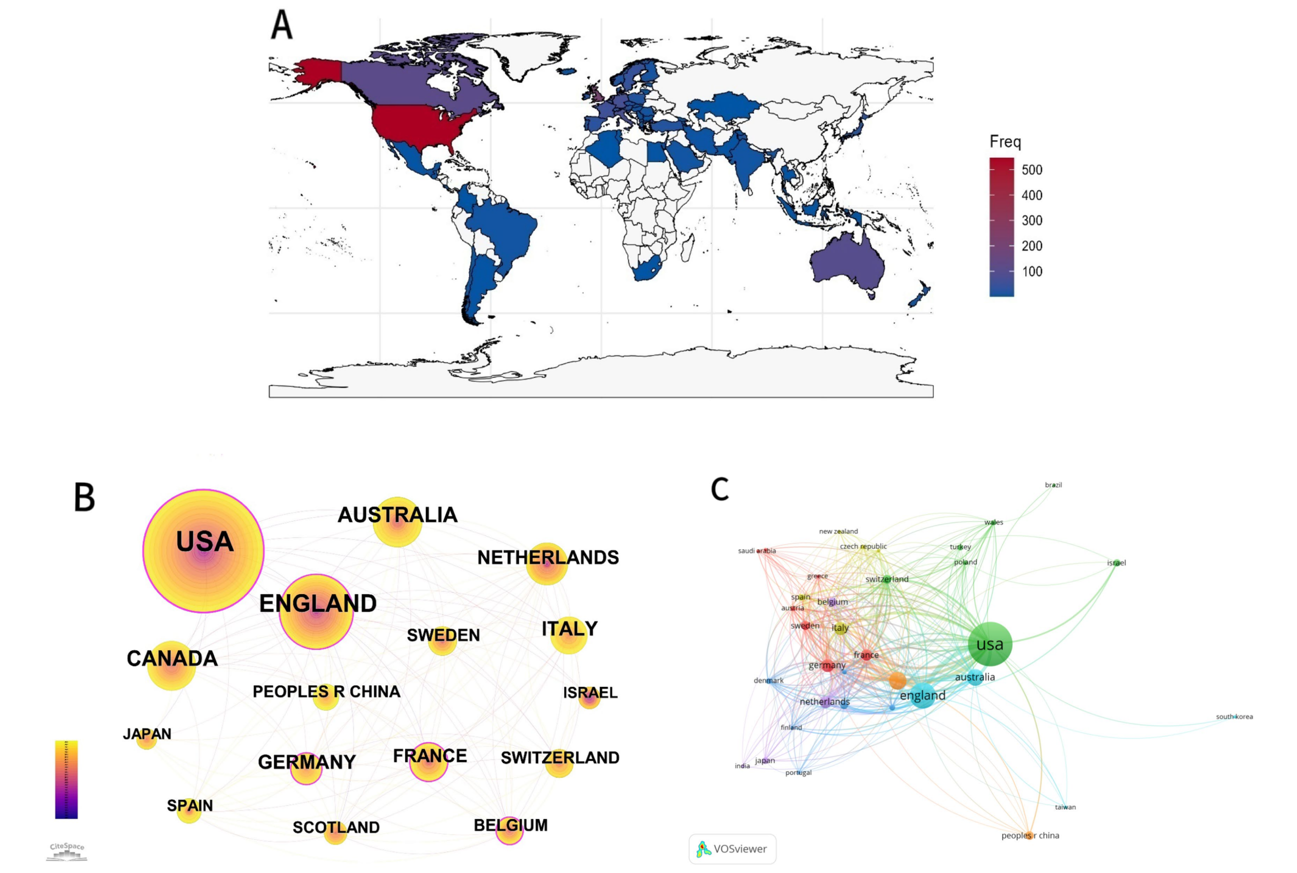


FIGURE 3
Contributions from various nations/regions to the broader autism phenotype. **(A)** Regional map of the broader autism phenotype studies' global spread. Different nations are represented by various hues depending on the number of articles published. More articles have been published when the color is redder; fewer articles have been published when the color is bluer. Because the map program cannot search for specific regions, we have provided the country in which the area is located directly. **(B)** A network diagram based on the study of country/regional collaboration and time-course-visualized using Cite Space. **(C)** A network diagram based on the study of country/regional collaboration and time-course-visualized using VOSviewer.

TABLE 1 The top twenty countries and centrality in the research field of BAP from 1994 to 2024.

Ranking	Countries	Frequency	Centrality
1	USA	546	0.32
2	ENGLAND	218	0.14
3	CANADA	115	0.07
4	AUSTRALIA	103	0.03
5	ITALY	73	0.04
6	GERMANY	65	0.14
7	NETHERLANDS	61	0.05
8	FRANCE	54	0.12
9	SWEDEN	38	0.03

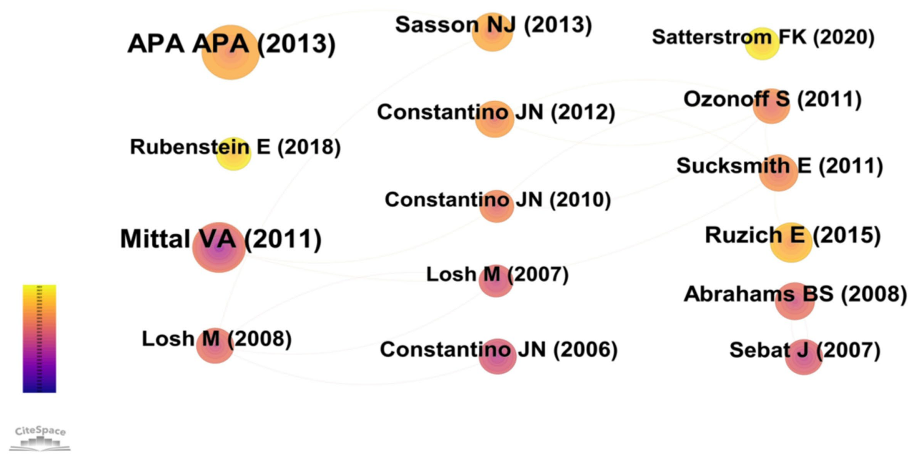


FIGURE 4
A dual-map overlay of the journals of BAP research.

3.4.1 Space

Upon selecting "institution" as the node type, we were able to generate an institutional distribution map (see Figure 5) that showed a combined network with 427 nodes and 2,539 links. Institutions are represented by nodes, and the relationships between them are shown by lines. Nodes that have published more papers are typically regarded as significant organizations that have a bigger influence on the advancement of a scientific field. A closer relationship is indicated by a thicker line connecting two nodes. Table 2 displays the top fifteen most productive institutions in this field of study. With 101 publications, the University of London is the most published institution overall. University of California System is ranked second with 79 publications, while King's College London is ranked third with 66 publications. The United States holds a prominent position in the field as evidenced by the majority of the top ten universities being situated within its borders. Furthermore, it is evident that there is a strong connection between the major institutions.

3.5 Authors analysis

We employed a 1-year time slice for our analysis and chose articles that were released between 1994 and 2024. We created a co-authorship network map in Cite.

Space by selecting authors as the node type. As shown in Figure 6, the combined network had 720 nodes and 1,343 links. Authors are represented by nodes, and author relationships are depicted by lines. In the collaboration Figure 6, the author's output is represented by the size of the node; the larger the node, the more articles the author has published. The degree of collaboration between them is indicated by the thickness of the line that separates them. Baron-cohen and Simon, the active authors from University of Cambridge's Departments of Experimental Psychology and Psychiatry, ranks first with 18 publications. Charman and Tony from University College London's psychology department comes in second with16, and Simon Baron-Cohen from the University of Cambridge's Autism Research Center comes in third with 15 publications. Table 3 displays the top ten authors in this field of study in terms of productivity. The writers are closely related to one another, and many of them have training in psychology.

A more accurate representation of the discipline's general trend can be seen in tables that examine the core authors' research directions and number of publications. Price's (1969) formula is applied here to determine a candidate for core authors. The equation is:

A research technique called co-citation is used to quantify the relationships between articles when one or more papers simultaneously

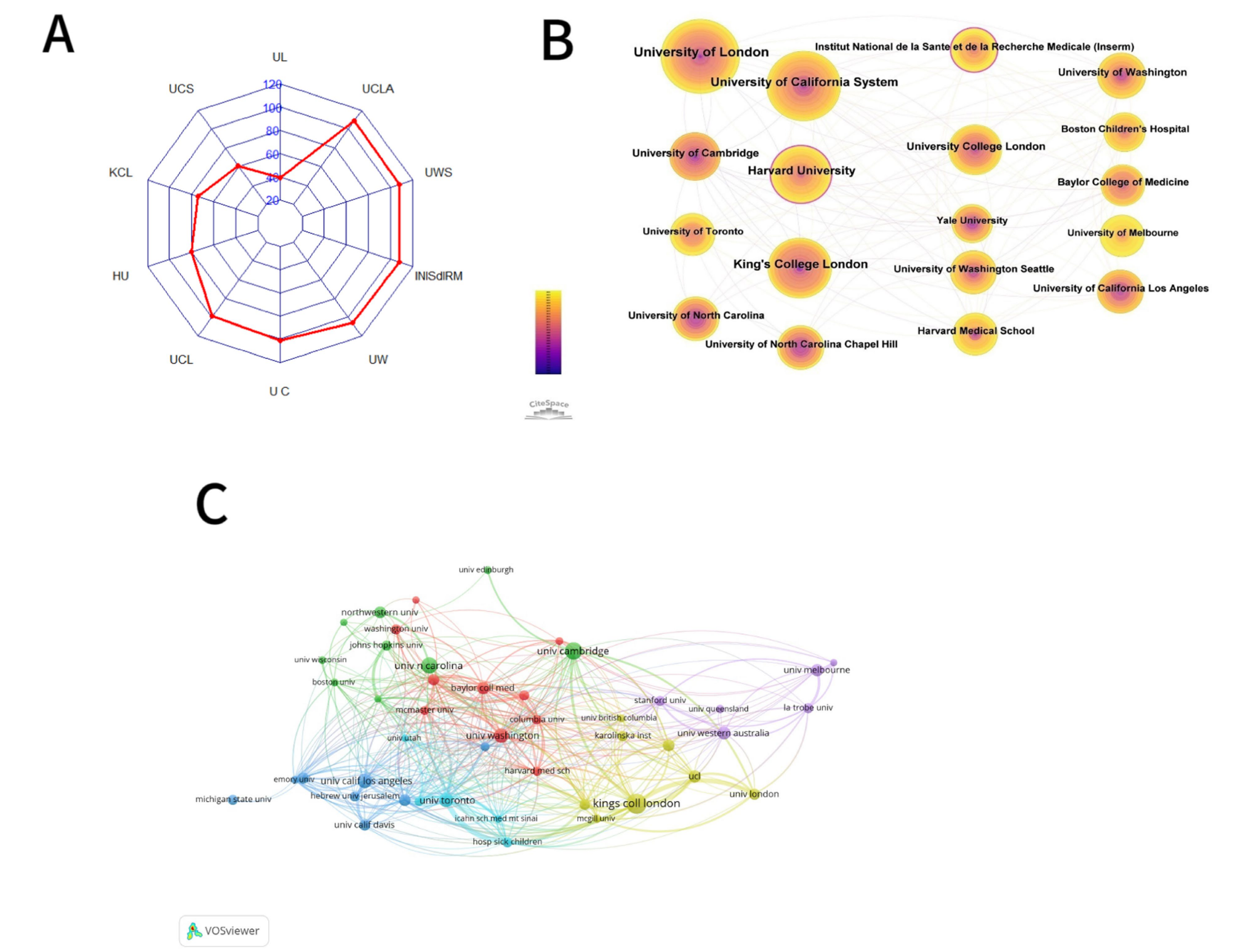


FIGURE 5
Visual analysis of institutes and funds. **(A)** Frequency radar map of the top ten funds. **(B)** Co-occurrence map of institutions of BAP from 2011 to 2024. **(C)** Visualization of the frequency of publications at each university.

TABLE 2 The top ten institutions and centrality in the research field of BAP from 2011 to 2024.

Ranking	Institutions	Frequency	Centrality	Half-life
1	University of London	101	0.09	17.5
2	University of California System	79	0.1	19.5
3	King's College London	66	0.05	17.5
4	Harvard University	60	0.12	18.5
5	University College London	41	0.05	13.5
6	University of Cambridge	39	0.06	17.5
7	University of Washington	34	0.04	13.5
8	Institut National de la Sante et de la RechercheMedicale (Inserm)	32	0.11	10.5
9	University of Washington Seattle	32	0.03	13.5
10	University of California Los Angeles	31	0.04	11.5

cite two or more articles, a relationship known as co-citation (Zhou et al., 2021). In CiteSpace, we chose references as the node type for our analysis, and we set a 1-year time slice for the articles' publication years between 1994 and 2024. This produced a co-citation network map, which, as shown in Figure 6, showed a merged network with 642 nodes and 2,631 links. Co-citations are represented by nodes, and the relationships between co-citations are shown by lines. Table 4 displays the ten most cited references in this study's film. Erratum: Ruzich et al. (2015)

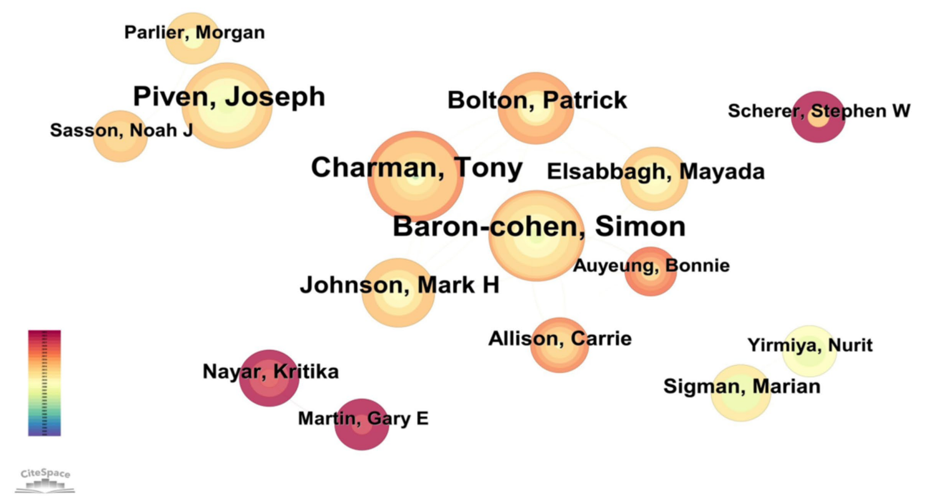


FIGURE 6
Co-occurrence map of author of BAP from 2011 to 2024.

TABLE 3 The top ten author and centrality in the research field of BAP from 2011 to 2024.

Ranking	Author	Frequency	Centrality	year	Half-Life
1	Losh, Molly	18	0.01	2007	13.5
2	Baron-cohen, Simon	16	0.01	2006	5.5
3	Charman, Tony	15	0.04	2003	9.5
4	Piven, Joseph	14	0.02	2007	2.5
5	Dell'osso, Liliana	10	0	2020	1.5
6	Bolton, Patrick	10	0.02	2009	5.5
7	Johnson, Mark H	9	0	2009	2.5
8	Carpita, Barbara	9	0	2020	2.5
9	Elsabbagh, Mayada	8	0	2009	2.5
10	Rutter, M	7	0	1995	2.5

most-cited work is Measuring Autistic Traits in the General Population: a systematic Review of the Autism-Spectrum Quotient (AQ) in a Nonclinical Population Sample of 6,900 Typical Adult Males and Females (Ruzich et al., 2015). The Broad Autism Phenotype Questionnaire: Prevalence and Diagnostic Classification was written by Sasson et al. (2013) came next, and followed Recurrence Risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium Study, authored by Ozonoff et al. (1994). It is evident that the majority of the co-citations discuss academic achievement, academic performance, and engagement.

3.6 Keywords analysis

Keywords are extremely concise and broadly defined, making it possible for readers to rapidly understand the main ideas of the study, This helps with hotspot analysis and development trend prediction. Understanding the core of a study requires an understanding of its keywords. We can find hotspots, examine research directions, and summarize the research topics in a particular field by examining keywords (Ma et al., 2021). A one-year time slice of papers published between 1994 and 2024 was chosen for our analysis, and keywords were selected as the CiteSpace node type. As seen in Figure 7, this produced a term co-occurrence map that showed a combined network with 584 nodes and 4,840 linkages. The study's most frequently occurring keywords were children (215), autism spectrum disorder (155), broad autism phenotype (140), phenotype (124), broad autism phenotype (98), parents (91), individuals (80), traits (73), and spectrum disorders (69).

In order to investigate the evolutionary trajectory and stage characteristics of the research field the timeline is predicated on the interactions and mutation relationships between keywords in a particular field. The co-occurrence map of keywords in the BAP timeline from 1994 to 2024 is shown in Figure 7 which also highlights hotspots and development directions in this field over time. The top 25 keywords with the most powerful citation bursts are displayed in Figure 8. The time intervals are shown by the blue line and the keyword outbreak time is indicated by the red line. The most recent terms to appear in the field of BAP research are epilepsy and autism spectrum disorder (after 2021). Earlier keywords in the field include family history autism spectrum disorders broader phenotype validity cognitive phenotype and toddlers (before 1994). Furthermore there is a lot of strength in genetics spectrum mental health and other factors.

10.3389/fpsyg.2024.1453140

TABLE 4 The top ten cited reference and centrality in the research field of BAP from 1994 to 2024.

Ranking	Cited reference	Year	Frequency	Centrality
1	Diagnostic and statistical manual of mental disorders, fifth edition, and the impact of events scale-revised	2013	67	0.01
2	Diagnostic and statistical manual of mental disorders	1994	64	0.01
3	Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females	2015	37	0.02
4	The Broad Autism Phenotype Questionnaire: Prevalence and Diagnostic Classification	2013	35	0.01
5	Advances in autism genetics: on the threshold of a new neurobiology	2008	32	0.02

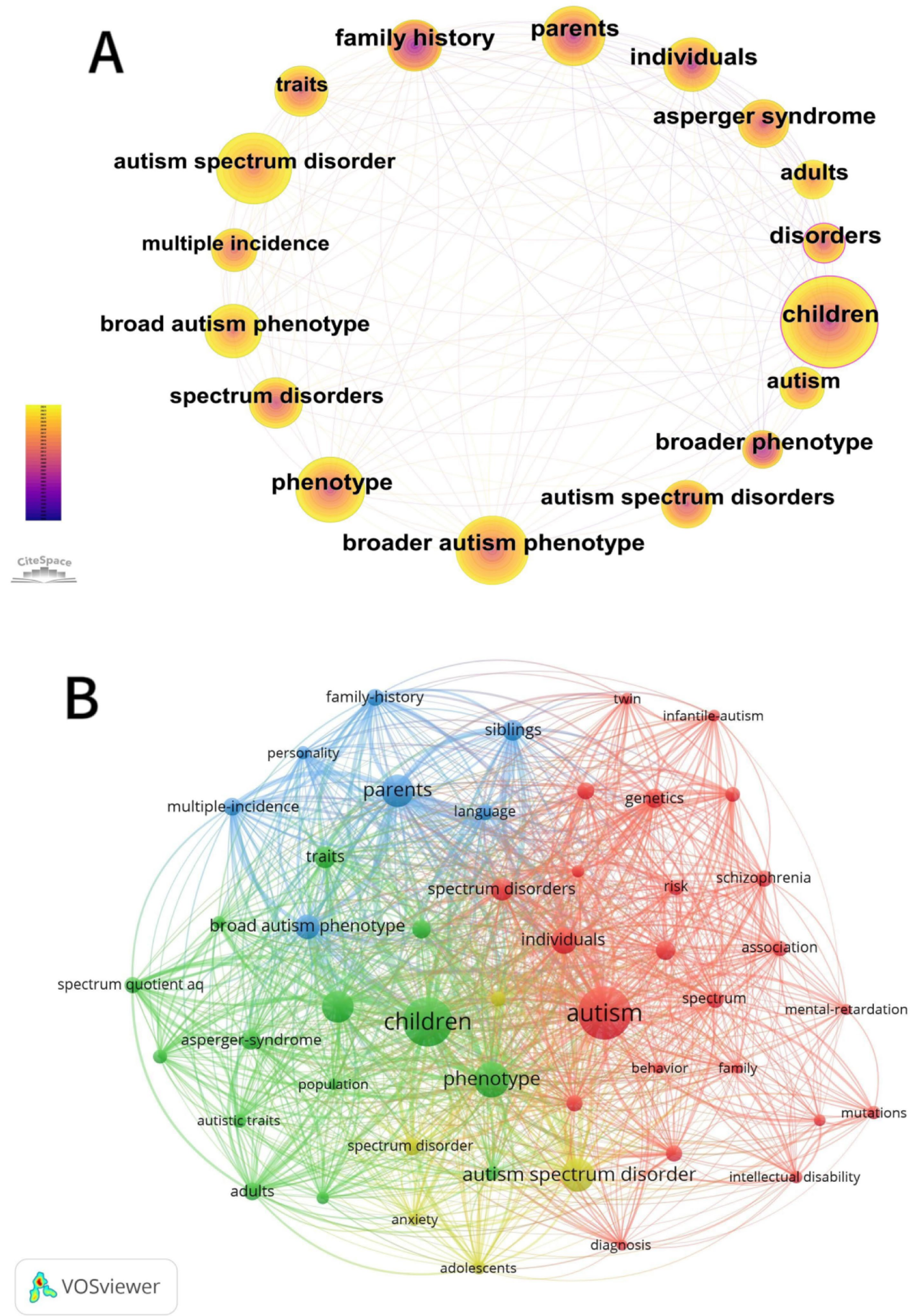


FIGURE 7 (Continued)

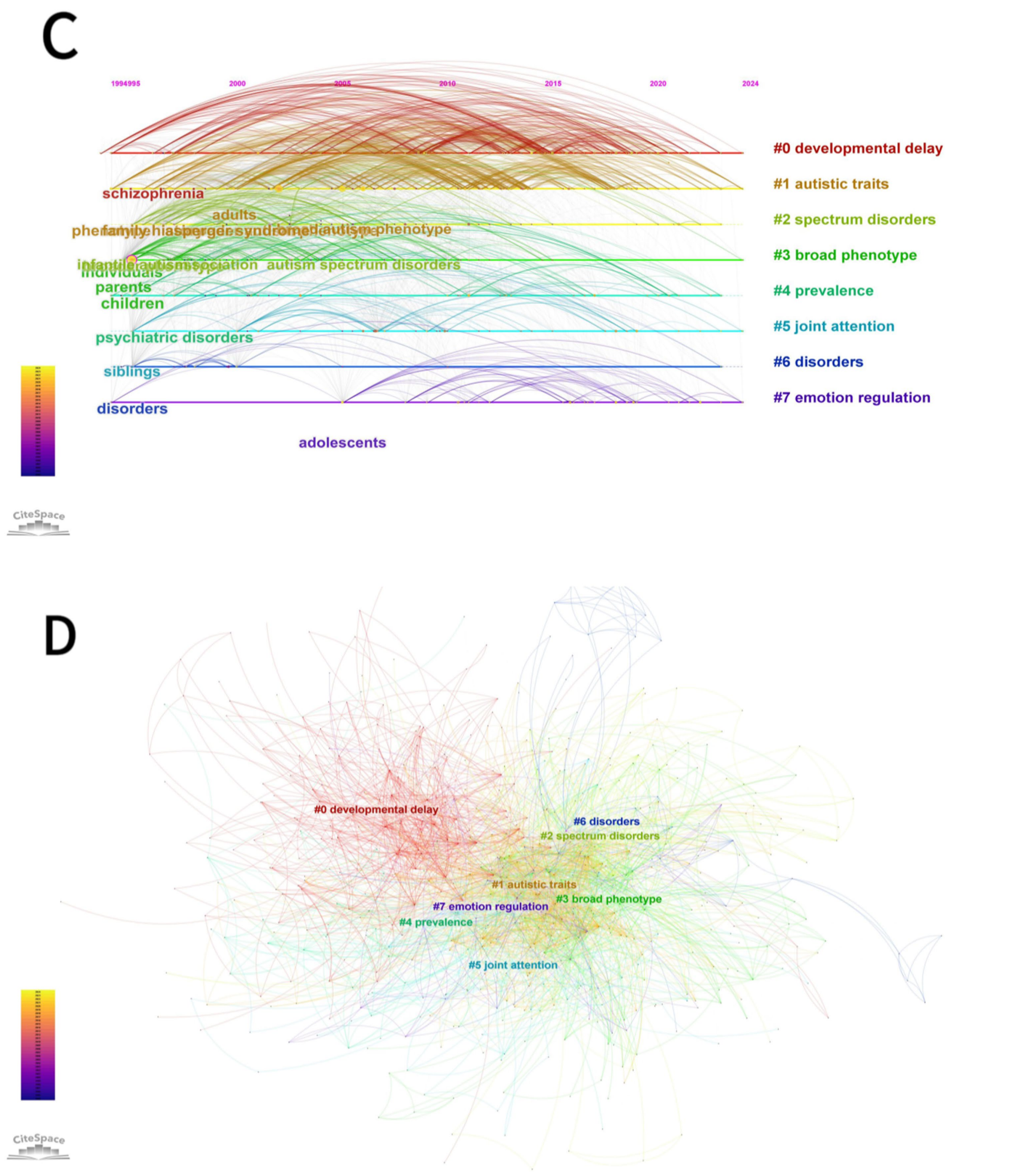


FIGURE 7
Visual map of keywords of BAP from 2011 to 2024. **(A)** The co-occurrence map of keywords. **(B)** Keywords co-occurrence analysis. Each keyword appeared at least 14 times. **(C)** Co-occurrence map of keywords timeline of BAP from 2011 to 2024. **(D)** The clustering map of keywords.

Future BAP trends may include the general population symptoms infants and adults

In the formula proposed by Donohue (1973):

$$T = \left[-1 + \sqrt{\left(\left(1 + 8I\right)\right)}\right] / 2.$$

T is the threshold of high-frequency keywords, and I is the total amount of keywords. High-frequency words can be easily located by computing the threshold (Kee and Tang, 2016). CiteSpace's metadata indicates that $I = 215$, from which $T \approx 20.24$ is computed, yielding a list of high-frequency keywords with a frequency more than 20.

4 Discussion

4.1 Global trends on FGCS

A quantitative method of examining academic literature, bibliometrics has been used in numerous fields of study to evaluate

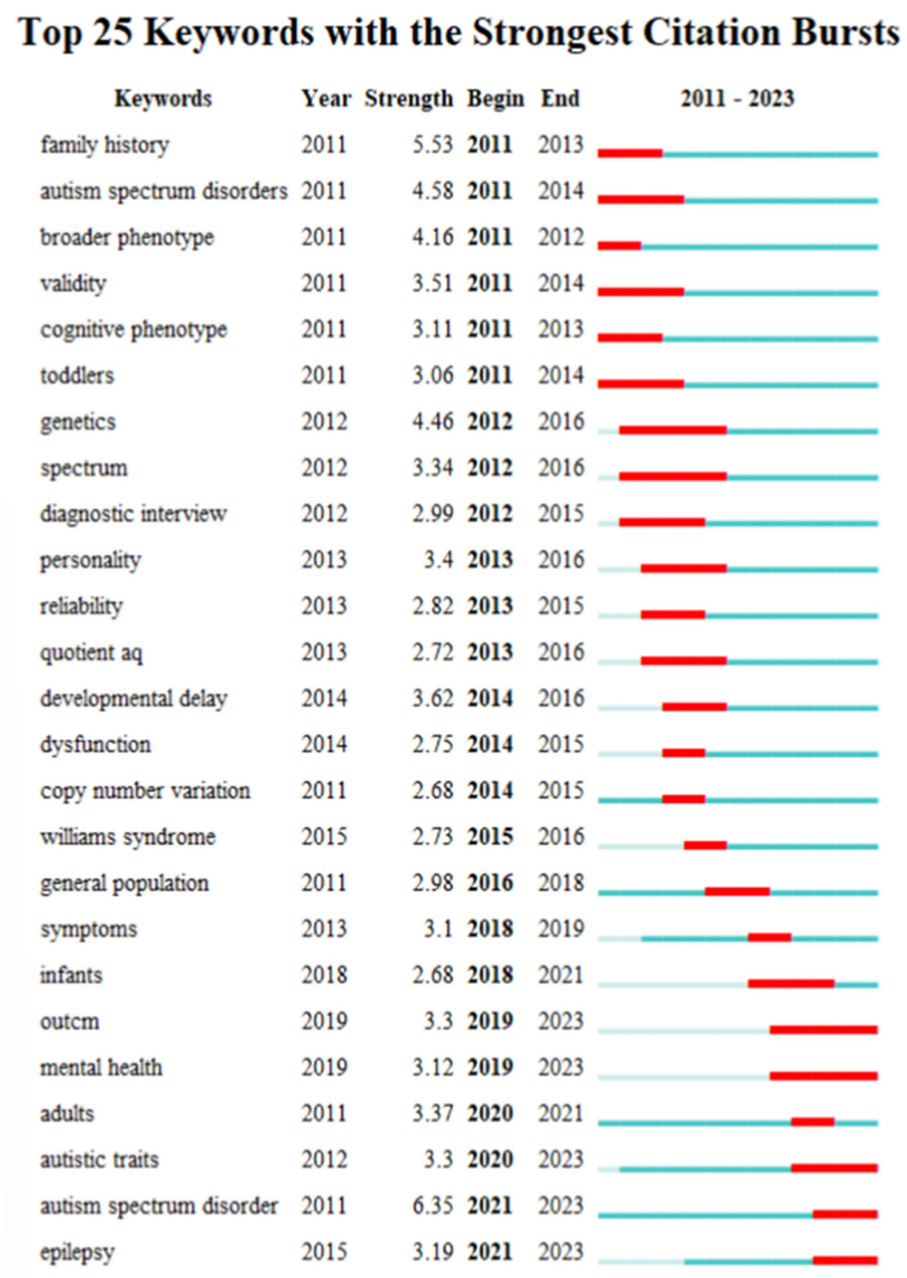


FIGURE 8
Visualization of top 25 keywords with the strongest citation bursts of BAP from 2011 to 2024.

the trends among authors, journals, institutions, nations, keywords, and other factors connected to specific publishing kinds. One visualization analytic tool called CiteSpace is frequently used to investigate research hotspots, research frontiers, knowledge bases, primary authors, and institutions in a field of study, and it is useful for forecasting a topic's future trajectory. It is also a useful option for determining over time the research trends and knowledge gaps in a field of study (Liang et al., 2017; Chen et al., 2022; Wu et al., 2021). Among the 1,075 literature we searched for, the earliest three articles were autistic spectrum disorder traits in children with attention deficit hyperactivity disorder, a cognitive endophenotype of autism in families with a multiple incidence and the relationship between the broader autism phenotype, child severity, and stress and depression in parents of children with autism spectrum disorders. "Autistic Spectrum Disorder Traits in Children with Attention Deficit Hyperactivity Disorder" researches whether autistic-like communication and social difficulties in children with ADHD are part of the broader ASD phenotype or are specific to ADHD. "A Cognitive Endophenotype of Autism in Families with Multiple Incidence" examines the neurocognitive endophenotype of autism, in families with multiple incidence autism and their findings do not confirm the hypotheses of weak central coherence or deficits in theory of mind as part of the broader endophenotype of autism. "The Relationship between the Broader Autism Phenotype, Child Severity, and Stress and Depression in Parents of Children with Autism Spectrum Disorders" examines the relationship between child symptom severity,

10.3389/fpsyg.2024.1453140

parent broader autism phenotype (BAP), and stress and depression in parents of children with ASD. The growth trend of BAP publications globally from 1994 to 2024 can be categorized into 3 phases. The first phase is, a slow upward trend from 1994 to 2006, the second phase is a downward trend from 2006 to 2013, and the third phase is a fluctuating trend from 2013 to today. During these 13 years, the highest number of publications was in 2018, 2021 and 2024 which both with 80 articles published. Nevertheless, BAP research lacks a description and assessment of the features of the literature, research orientations, research depth, and research hotspots.

The total number of articles published in the last 13 years has gone through some ups and downs, but has generally shown an upward trend to reach a total of 1,075. A paper with 100 or more citations is typically regarded as a "classic" or possibly even a seminal paper in the field of study. Generally speaking, the highest number of citations per year for a paper usually occurs between 3 and 10 years after publication (Xiong et al., 2021). The highest number of citations in this field of study is Genetic Heritability and Shared Environmental Factors Among Twin Pairs with Autism with 1,541 citations and Recurrence Risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium Study 1,130 citations. Through analysis of the literature, we find that susceptibility to ASD has moderate genetic heritability and a substantial shared twin environmental component and the sibling recurrence rate of ASD is higher than suggested by previous estimates (Hallmayer et al., 1994; Ozonoff et al., 1994). Analyses of the papers indicate that research in this area has focused on ASD, suggesting that the research on ASD is extremely essential when it comes to BAP.

4.2 Research focus and hotspots on BAP

Analysis of the results of high-frequency and strongest burst keywords showed that the research focus and hotspots of BAP has changed over time. Since 1994, people have paid more attention to the study of BAP in families, emphasizing the relationship between symptoms of parents and children. Any child intervention plan must take into account the parent and family factors. Understanding that some family members have BAP traits should be a part of overall family considerations when it comes to ASD, especially in families with multiple affected children. A thorough family history taken before treatment begins can also be useful in identifying whether a multiplex family has a higher likelihood of BAP expression and in clinical settings when recommending the best course of action for a particular child (Gerdts et al., 2013). The presence of BAP in parents of ASD children was found to be substantially linked to the child's phenotypic class, which included mild language and motor delays, average nonverbal abilities, and more co-occurring conditions like anxiety, depression, and sleep issues (Rubenstein et al., 2019).

Meanwhile, BAP research also focuses on cognitive aspects. More strongly than other neuropsychological characteristics of autism, social-cognitive deficits distinguish parents with the "broad autism phenotype" from parents without it. This suggests that this domain may be especially useful for identifying genetic and brain processes linked to the phenotype (Sasson et al., 2013). And cognitive barriers are also age-related. When people with the BAP grew older, their subjective cognitive impairment increased even though they did not experience any other cognitive decline or accelerated memory loss (Caselli et al., 2018). According to a study, about one in two adults and adolescents with the BAP showed a decreased capacity to recognize and understand social faux pas and there were no gender differences in advanced ToM abilities. A key endophenotype in a subset of people with the BAP may be impaired advanced ToM abilities (Green et al., 2020).

The study of ASD has been a hot topic and the latest trend in the study of BAP since 1994. Meanwhile, research on AT has also become a hot topic in BAP in recent years. In a study, researchers used AQ (Autism Spectrum Quotient) to evaluate autistic traits in participants and found that a significant proportion of fathers and mothers of ASD children belonged to the high-scoring group, indicating that BAP was not sensitively detected among parents of ASD children by paying attention to detail (Bora et al., 2017). Studies on the BAP could help those who live with family members who are autistic by offering more support and direction. For instance, fathers of autistic children who exhibit the BAP could be given tips on how to handle and strengthen their relationships with peers and other family members (Sucksmith et al., 1994).

4.3 Study strengths and limitations

Through the creation of knowledge maps and timelines, bibliometric analysis can perform visual analysis on bibliographic data, giving academics a better understanding of the connections and development trends among publications. This study examines the global trends and future prospects in BAP research during the previous 13 years through a projected bibliometric analysis of BAP. Additionally, the search was not restricted to a single academic publication but rather used WoSCC as a database to yield rich data thanks to the precise and thorough literature search criteria. The bibliometric analysis of this study also included popular subject categories, keyword analysis, nation and institutional productivity, annual publishing output, the most cited reference, country, institution, and journal distribution.

Ultimately, a few restrictions need to be taken into account. First, we did not use any other electronic databases like PubMed or Embase; instead, we only used the WoSCC data that was available. Consequently, there could be a delay in the papers gathered from the WoSCC database, which would introduce bias into the study's citation counts and H-indexes. Secondly, the citation analysis could be biased by nature. Citation rates differ according on specialization and the size of the field of study. There are typically more classical references in more popular scientific subjects. Furthermore, authors are more likely to cite their own works, and papers written in English are more likely to be cited in general. Thirdly, the legitimacy or scientific rigor of publications is not adequately taken into account by bibliometric approaches. Publications with a high citation count may not always be of excellent scientific caliber. Finally, since we were just using BAP as a search keyword, it's possible that some articles were missed. Also, the scope of our analysis was limited to the last 13 years, leaving out older research on BAP, which may have contributed to its incomplete representation.

In conclusion, despite these drawbacks, we still think that scholars can gain a deeper and more thorough understanding of the development trends and research hotspots of publications pertaining to BAP by utilizing CiteSpace for visual statistical analysis.

5 Conclusion

This study gathered a significant amount of reliable data by conducting a bibliometric analysis of the BAP literature published in the previous 12 years. It also summarized previous research, replicated the BAP research process, investigated global research hotspots, and evaluated the boundaries of BAP development in the private sector. The study's findings demonstrate that, over the previous 20 years, there has been a roughly global increase in the total number of papers pertaining to BAP, suggesting that BAP has attracted a lot of attention. Through the analysis of 1,075 articles, we discovered that this field of research has experienced significant growth after 2017, suggesting that it is still a young field with a wealth of untapped research opportunities. The two most popular research areas are sociology and psychology. More than any other nation, the United States is the primary research force in this area. The remarkable thing about the BAP research field is how many countries, institutions, and authors have joined forces to collaborate extensively. Our analysis of the clustering results and keyword frequency in the scholarly literature reveals that the hotspots and focus of BAP research have evolved over time. Some research hotspots are summarized as follows:

(1) Numerous phenotypic-focused research hotspots emerged in 1994, examining the relationship between cognitive phenotype, broader phenotype, and autism spectrum disorders. Researchers have increasingly increased the scope of BAP research to include the general population such as infants and adults as well as paying greater attention to the mental health concerns that BAP experience in recent years.
(2) The BAP's academic issues have been a significant area of study for almost 10 years, and the obstacles and problems the BAP faces in the academic setting have garnered a lot of attention.
(3) Research on the specific symptoms of BAP (such developmental delay) and how it relates to autistic traits and autism spectrum disorder has become increasingly popular as the field grows.
(4) Research trends in the future may concentrate on social support, cultural identity, and other topics in addition to the previously mentioned factors.

In summary, this study presents a historical, forward-looking perspective and provides a comprehensive and reliable body of research that provides an accurate overview of the global landscape of BAP research. Although these records do not.

The seven citations with the most intense bursts (sorted by year of bursts onset). The blue line represents the time when the citation appeared, and the red line is the time when the bursts happened. Encompass every single publication on BAP, they do provide a significant and representative sample that can offer valuable insights into the challenges and opportunities faced by this population.

Overall, this study sheds information on important topics including familial inheritance and particular BAP symptom presentations. Further thorough and in-depth research on BAP is needed.

Author contributions

F-QQ: Writing – original draft, Writing – review & editing. S-NL: Investigation, Writing – original draft, Writing – review & editing. T-TD: Conceptualization, Software, Writing – review & editing. W-MC: Data curation, Writing – review & editing. Y-YS: Software, Validation, Writing – original draft. XQ: Data curation, Methodology, Writing – original draft. Y-JD: Methodology, Software, Writing – original draft. LW: Investigation, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was supported by the construction of the psychological counseling case database for the "Unveiling Project" of the "Young Talent Lift Project" of Shandong Psychological Society in 2024 (Project No.: SDPSYESS2024001).

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

References

Bailey, A. J., Couteur, A. L., Gottesman, I. I., Bolton, P. F., and Rutter, M. L. (1995). Autism as a strongly genetic disorder: evidencefrom a british twin study. *Psychol. Med.* 25, 63–77. doi: 10.1017/S0033291700028099

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The autism-spectrum quotient (aq): evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17. doi: 10.1023/A:1005653411471

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., et al. (2010). A case-control family history study of autism. *J. Child Psychol. Psychiatry* 35, 877–900. doi: 10.1111/j.1469-7610.1994.tb02300.x

Bora, E., Aydın, A., Saraç, T., Kadak, M. T., and Köse, S. (2017). Heterogeneity of subclinical autistic traits among parents of children with autism spectrum disorder: identifying the broader autism phenotype with a data-driven method. *Autism Res.* 10, 321–326. doi: 10.1002/aur.1661

Caselli, R. J., Blake, T., Amylou, C., Dona, E. C., and Woodruff, B. K. (2018). Subjective cognitive impairment and the broad autism phenotype. *Alzheimer Dis. Assoc. Disord.* 32, 284–290. doi: 10.1097/WAD.0000000000000273

Chen, C. (2004). Searching for intellectual turning points: progressive knowledge domain visualization. Proc. natl. acad. Sei. Usa, 101 (suppl), 5303-5310. *Proc. Natl. Acad. Sci.* 101, 5303–5310. doi: 10.1073/pnas.0307513100

Chen, B., Shin, S., Wu, M., and Liu, Z. (2022). Visualizing the knowledge domain in health education: a Scientometric analysis based on cite space. *Int. J. Environ. Res. Public Health* 19:6440. doi: 10.3390/ijerph19116440

Donohue, J. C. (1973). *Understanding scientific literatures: A bibliometric approach*. MIT Press.

Folstein, S., and Rutter, M. (1977). Genetic influences and infantile autism. *Nature* 265, 726–728. doi: 10.1038/265726a0

Gao, M., Zhang, H., Gao, Z., Sun, Y., Wang, J., Wei, F., et al. (2022). Global hotspots and prospects of perimenopausal depression: a bibliometric analysis via cite space. *Front. Psychol.* 13:968629. doi: 10.3389/fpsyt.2022.968629

Gerdts, J. A., Bernier, R., Dawson, G., and Estes, A. (2013). The broader autism phenotype in simplex and multiplex families. *J. Autism Dev. Disord.* 43, 1597–1605. doi: 10.1007/s10803-012-1706-6

Green, C. C., Brown, N. J., Yap, V. M. Z., Scheffer, I. E., and Wilson, S. J. (2020). Cognitive processes predicting advanced theory of mind in the broader autism phenotype. *Autism Res.* 13, 921–934. doi: 10.1002/aur.2209

Hallmayer, J. F., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., et al. (1994). Genetic heritability and shared environmental factors among twin pairs with autism. *Arch. Gen. Psychiatry* 68, 1095–1102. doi: 10.1001/archgenpsychiatry.1994.76

Hurley, R. S. E., Losh, M., Parlier, M., Reznick, J. S., and Piven, J. (2007). The broad autism phenotype questionnaire. *J. Autism Dev. Disord.* 37, 1679–1690. doi: 10.1007/s10803-006-0299-3

Kanner, L. (1943). Autistic disturbances of affective contact. *Acta Paedopsychiatr.* 35, 100–136.

Kanne, S. M., Wang, J., and Christ, S. E. (2012). The subthreshold autism trait questionnaire (satq): development of a brief self-report measure of subthreshold autism traits. *J. Autism Dev. Disord.* 42, 769–780. doi: 10.1007/s10803-011-1308-8

Kee, H. L., and Tang, H. (2016). Domestic value added in exports: Theory and firm evidence from China. *Am. Econ. Rev.* 106, 1402–1436. doi: 10.1257/aer.20131687

Lainhart, J. E., Ozonoff, S., Coon, H., Krasny, L., Dinh, E., Nice, J., et al. (2002). Autism, regression, and the broader autism phenotype. *Am. J. Med. Genet.* 113, 231–237. doi: 10.1002/ajmg.10615

Liang, Y. D., Li, Y., Zhao, J., Wang, X. Y., Zhu, H. Z., and Chen, X. H. (2017). Study of acupuncture for low back pain in recent 20 years: a bibliometric analysis via cite space. *J. Pain Res.* 10, 951–964. doi: 10.2147/JPR.S132808

Liu, S., Sun, Y. P., Gao, X. L., and Sui, Y. (2019). Knowledge domain and emerging trends in Alzheimer's disease: a scientometric review based on cite space analysis. *Neural Regen. Res.* 14, 1643–1650. doi: 10.4103/1673-5374.255995

Luo, H., Cai, Z., Huang, Y., Song, J., Ma, Q., Yang, X., et al. (2021). Study on pain catastrophizing from 2010 to 2020: a bibliometric analysis via citespace. *Front. Psychol.* 12:759347. doi: 10.3389/fpsyg.2021.759347

Ma, D., Yang, B., Guan, B., Song, L., Liu, Q., Fan, Y., et al. (2021). A bibliometric analysis of Pyroptosis from 2001 to 2021. *Front. Immunol.* 12:731933. doi: 10.3389/fimmu.2021.731933

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., et al. (1994). Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. *Pediatric* 128, e488–e495. doi: 10.1542/peds.2010-2825

Price, D. J. de S. (1969). Networks of scientific papers. *Science*, 166, 913–914. doi: 10.1126/science.166.3906.913

Rubenstein, E., Wiggins, L. D., Schieve, L. A., Bradley, C., DiGuiseppi, C., Moody, E., et al. (2019). Associations between parental broader autism phenotype and child autism spectrum disorder phenotype in the study to explore early development. *SAGE* 23, 436–448. doi: 10.1177/1362361317753563

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., et al. (2015). Measuring autistic traits in the general population: a systematic review of the autism-spectrum quotient (aq) in a nonclinical population sample of 6, 900 typical adult males and females. *Mol. Autism.* 6, 1–12. doi: 10.1186/2040-2392-6-2

Sasson, N., Lam, K. S., Childress, D., Parlier, M., Daniels, J. L., and Piven, J. (2013). The broad autism phenotype questionnaire: prevalence and diagnostic classification. *Autism Res.* 6, 134–143. doi: 10.1002/aur.1272

Sasson, N. J., Nowlin, R. B., and Pinkham, A. E. (2013). Social cognition, social skill, and the broad autism phenotype. *Autism* 17, 655–667. doi: 10.1177/1362361312455704

Sucksmith, E., Roth, I., and Hoekstra, R. A. (1994). Autistic traits below the clinical threshold: re-examining the broader autism phenotype in the 21st century. *Neuropsychol. Rev.* 21, 360–389. doi: 10.1007/s11065-011-9183-9

Sung, Y. J., Dawson, G., Munson, J., Estes, A., Schellenberg, G. D., and Wijsman, E. M. (2005). Genetic investigation of quantitative traits related to autism: use of multivariate polygenic models with ascertainment adjustment. *Am. J. Hum. Genet.* 76, 68–81. doi: 10.1086/426951

Wheelwright, S., Auyeung, B., Allison, C., and Baron-Cohen, S. (2010). Defining the broader, medium and narrow autism phenotype among parents using the autism spectrum quotient (aq). *Mol. Autism.* 1, 1–9. doi: 10.1186/2040-2392-1-10

Wu, M. Q., Wu, D. Q., Hu, C. P., and Iao, L. S. (2021). Studies on children with developmental coordination disorder in the past 20 years: a bibliometric analysis. *Front. Psych.* 12:776883. doi: 10.3389/fpsyt.2021.776883

Xiong, H. Y., Zhang, Z. J., and Wang, X. Q. (2021). Bibliometric analysis of research on the comorbidity of pain and inflammation. *Pain Res. Manag.* 2021:211. doi: 10.1155/2021/6655211

Zhou, Q., Kong, H. B., He, B. M., and Zhou, S. Y. (2021). Bibliometric analysis of bronchopulmonary dysplasia in extremely premature infants in the web of science database using citespace software. *Front. Pediatr.* 9:705033. doi: 10.3389/FPED.2021.705033

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 03 January 2025
DOI 10.3389/fpsyg.2024.1455941

Check for updates

OPEN ACCESS

EDITED BY
Douglas Frye,
University of Pennsylvania, United States

REVIEWED BY
Gary Morgan,
Fundació per a la Universitat Oberta de Catalunya, Spain
Elisabet Serrat,
University of Girona, Spain
Ira Andrew Noveck,
Centre National de la Recherche Scientifique (CNRS), France
Kevin John Riggs,
University of Hull, United Kingdom

*CORRESPONDENCE
Jill G. de Villiers
✉ jdevilli@smith.edu

†These authors have contributed equally to this work and share first authorship

RECEIVED 27 June 2024
ACCEPTED 09 December 2024
PUBLISHED 03 January 2025

CITATION
de Villiers JG and de Villiers P (2025) The contributions of language and inhibitory control to false belief reasoning over time. *Front. Psychol.* 15:1455941.
doi: 10.3389/fpsyg.2024.1455941

The contributions of language and inhibitory control to false belief reasoning over time

Jill G. de Villiers*† and Peter de Villiers†

Department of Psychology, Smith College, Northampton, MA, United States

Introduction: The role of language in false belief reasoning has been much debated for twenty-five years or more, especially the relative contributions of general language development, complement syntax, vocabulary, and executive function. However, the empirical studies so far have fallen short, in that they generally have too few participants for adequate statistical modeling; they do not include control variables; or they are cross-sectional rather than longitudinal, making inferences about causal direction much more tenuous.

Methods: The present study considers the role of these different variables in the development of false belief reasoning over several months of testing, with 258 children aged three to five years. The children are also from under-resourced communities, broadening the populations that generally contribute such data.

Results: A cross-sectional and a longitudinal regression analysis reveals the contribution of each variable to the children's success on the false belief measures. Finally, a structural equation model tests the relative contribution of the different potential factors over time, how they interact, and change. The model is an excellent fit to the data. Inhibitory control, complement comprehension and vocabulary all have effects on false belief reasoning at the first time point (T1). However, at T3, the major proximal contribution is the child's comprehension of complements, though the longitudinal pathways of vocabulary and inhibitory control also pave the way.

Discussion: Our data confirm the specific contribution of complement syntax but also makes clear, as do training studies, that a certain amount of preparedness in vocabulary and in executive function skills is also necessary.

KEYWORDS

theory-of-mind, complements, structural equation, language, inhibitory control, longitudinal

1 Introduction

1.1 The course of early theory of mind

Theory of mind, defined as the understanding that others' mental states may be different from one's own, has been studied exhaustively for four decades now, and there remain many debates about the influences on the developmental path children take. The general developmental course itself is not disputed, though the extent of cross-cultural variation is still not established. The stepwise development defined by Wellman and Liu (2004) in terms of definable and well-matched experimental tasks is generally accepted, though there is still uncertainty about the causal necessity of each stage. The first step in their developmental sequence, at around age 18 months, is understanding others' desires, that is, recognizing that two people may want different things than the child does (Repacholi and Gopnik, 1997).

Following this, the child comes to understand that people will act according to their beliefs, even if different from the child's. Understanding that "seeing leads to knowing," or knowledge access, is slightly later. Realizing that others can have false beliefs, when the child has evidence to the contrary, comes in around age four or five years. Finally, understanding "hidden emotions" is found to be a still later development (Wellman et al., 2011). However, the sequence of middle steps may be different for Chinese children, where understanding different beliefs may follow knowledge access (Wellman et al., 2006; Yu and Wellman, 2024). Still further theory of mind skills, such as second order false belief reasoning (Miller, 2009) and understanding sarcasm and irony (Filippova and Astington, 2010; Winner, 1997) do not develop until early school age.

Digging deeper, it is clear there are also important developments before age two. In the first year of life, children are sensitive to others' intentions and goals, and differentiate between intentional and unintentional actions (Woodward, 2009). When learning their first language, children make use of shared attention with a caregiver to establish the referents of early words (Baldwin, 1995; Tomasello, 1999). Furthermore, infants in the second year of life can track which objects another person has seen before, and will show that person toys that they have not seen, or point out to them where an object was hidden (Tomasello, 1999).

Why has false belief reasoning occupied center stage in the literature about theory of mind? Philosophers and psychologists have considered the achievement of false belief understanding to be a watershed not only in early child development (Dennett, 1978; Wellman et al., 2001) but also in assessing the theory of mind capabilities of other species (Call and Tomasello, 2008; Premack and Woodruff, 1978) and recently, of A.I. (Strachan et al., 2024) In a false-belief task, you must hold in memory alternative perspectives within a given context (e.g., where a target object actually is and where a character thinks it is). You then must also inhibit the prepotent response of selecting a perspective that is congruent with your own (i.e., selecting where you know the target object to be) (Leslie et al., 2005). To understand another's false beliefs is to entertain ideas inconsistent with the reality you know, and to reason from that representation about either the cause or the consequence of that belief. Westra and Nagel (2021) refer to this as *non-factive* reasoning, in contrast to the *factive* reasoning that is entailed in recognizing what someone knows. In recognizing knowledge, you pay attention to whether the person has perceptual access to the thing or event in question, and since you share the perception, there is no contradiction in perspective. We know that quite young children are active observers of attention and perceptual access, but do not seem attentive to mistaken beliefs.

Since the work of Wellman and Liu (2004), there has been an explosion of research on infant theory of mind, with some researchers claiming that given the right tasks, infants can be shown to recognize that another character has a false belief (Baillargeon et al., 2010; Scott and Baillargeon, 2017). The theory of mind tasks used with infants are called implicit tasks, because no behavioral decision is required from the infant, that is, there is no step of response selection. The first type of task uses the length of gaze as the measure. Infants in the second year of life or even younger have been shown to gaze for a more protracted time at events in which a human character acts in a way contrary to expectation, namely, a way that is not in keeping with the belief they should have formed (Onishi and Baillargeon, 2005). For example, in a version of an unseen displacement task, they are surprised when the character being observed goes to a location where an object really is, when the character did not see it move there. In a second task type, infants look expectantly at a location where a character should go, based on where that character falsely believes something to be (Southgate et al., 2007). In a third design, very young children come to the assistance of another individual specifically if that person was not witness to how something works (Buttelmann et al., 2009). Clearly, success at a young age on implicit tasks calls into question the necessity of language as an aid to reasoning.

The age gap between these tasks and explicit false belief mastery around age four requires explanation. On the account provided by several researchers (e.g., Baillargeon et al., 2010), the gap is a result of the task demands of explicit tasks, namely, response selection sets the bar too high for infants. On this account infants have implicit understanding of others' beliefs, as rich as that of older children or adults (Baillargeon et al., 2018). However, others ask: Are the infants in these studies acting on the basis of a belief attribution, or something simpler (Southgate, 2013)? The most reductive explanation is that the infant is responding to some accidental but correlated feature of the set-up (Heyes, 2014). Some have found a lack of connection between the looking tasks and the helping tasks (e.g., Poulin-Dubois and Yott, 2018), arguing against abstract mental states underlying both. Other theorists suggest that the child is responsive to a behavioral rule, such as "people go to where they last saw something" (Perner and Ruffman, 2005). Southgate and Vernetti (2014) suggested that infants may be able to follow an agent's point of view at ages as young as 6 months, but the difference is that they do not yet contrast it with their own. In a widely-cited account, Apperly and Butterfill (2009) proposed that infants may "register" another's belief rather than represent it. Registration allows infants to trace and track another's belief but in a limited way that betrays it as different from mature representation, for example, if the belief being tracked has relationships to other mental states, or involves subtle properties not easily captured by vision. As one example, registration may be restricted to tracking expected location rather than how a person construes an object (see also Low, 2010). On Apperly and Butterfill's two-systems account, the registration that infants use gives way to real representation in older children.

Unfortunately, these fascinating ideas remain in doubt due to a replicability crisis. Many failures to replicate the findings have been reported across developmental labs, and the difficulty is that failures do not get published (Kulke and Rakoczy, 2018; Kulke et al., 2018; Rakoczy, 2012; but see Baillargeon et al., 2018). Brooks and Meltzoff (2015) found continuity in a small group of nineteen children from the first implicit understanding in anticipatory looking in infancy at 10.5 months through to explicit false belief development at age 4.5 years. However, implicit theory of mind tasks used beyond infancy do not show a consistent relationship to explicit tasks with the same preschool children (Low, 2010; Grosse Wiesmann et al., 2017). Nevertheless, the putative gap between infant and four-year-old theory of mind needs explanation and a popular suggestion is that the child needs to develop a mature executive function before success on explicit false belief tasks, because of the cognitive demands of the task. We will refer to this as the *executive function hypothesis*.

1.2 Executive functions and theory of mind

Executive function development is on a time course that has some correspondence with theory of mind changes, with some significant maturation happening around four or five years of age. Two aspects in particular have been highlighted as potential determinants of the change in false belief reasoning development in previous research. One is *working memory* (Brandt et al., 2023; Carlson et al., 2002), and the second is *inhibitory control* (Carlson et al., 2015; Carlson and Moses, 2001; Carlson et al., 2002). In an explicit false-belief task, it is necessary to hold alternatives in memory (e.g., what an object actually is and what someone thinks it is). The child also has to inhibit the alternative that is in keeping with their own beliefs, and to resist that lure of reality. Inhibitory control would also seem to be prerequisite for the step of inhibiting the reality response. In fact, in an important meta-analysis of previous work on executive function and false belief reasoning, Devine and Hughes (2014) report that inhibitory control is a reliable predictor of false belief understanding in preschoolers. Nevertheless, is it important to consider the role of executive function in conjunction with the role of another significant skill developing over the preschool years, namely language.

1.3 Language and theory of mind

Alternative proposals have highlighted the role of language in the process of theory of mind development (de Villiers and de Villiers, 2009), and some theorists argue that language is an alternative to bridge the gap from infant implicit understanding to explicit false belief reasoning (Apperly and Butterfill, 2009). Studies have found that the child's own language appears to be a key factor in mastery of explicit false belief understanding in young children (Astington and Baird, 2005; San Juan and Astington, 2012). However, is it learning mental state vocabulary (Shatz, 1994), engaging in rich discourse (Harris et al., 2005; Nelson, 2005), or acquiring grammatical structures that contribute to general reasoning (Farrar et al., 2017; Ruffman et al., 2002). In particular, are specific syntactic achievements necessary (de Villiers and de Villiers, 2000, 2014)?

1.3.1 Sentential complements

Explicit false belief reasoning requires minimally two processes:

(a) The representation of alternatives, one of which is false.
(b) The ability to inhibit a prepotent response in response selection.

The representation of false alternatives is much more easily done with language than with images or even words (Nordmeyer and de Villiers, 2019). Neither images nor words can be false, as they are not propositions in themselves. Of course, one can form different images representing different perspectives, but not one that is false, or negative, without something extra.

> "Consider, for example, negation. It's easy to tell somebody that it's not going to rain. Try drawing them a picture of it's not going to rain … Think about trying to draw a picture of "there's not a giraffe standing beside me" (Fodor, 1994).

This is trivially easy to do, unless one needs to recognize what it is a picture of. Language itself has rich ways of representing not only different perspectives, for example through deixis (de Villiers, 2018), but it can also represent negation and other logical terms that images do not. The unique part about certain embedded sentences, or sentential complements, is that they can represent a false proposition inside a true sentence.

de Villiers (2007) has argued that complements constitute an example of language-as-cognitive tool that has special utility in representing the states of others' minds. The complement in (1) is distinct from the adjunct clause in (2), because the embedded proposition in (1) can be false:

(1) Miriam said *that she baked the bread*.
(2) Miriam relaxed *after she baked the bread*.

The complement structure only occurs under communication and mental state verbs. Thus, complements can express mistakes and lies, and with mental state verbs like *think* or *believe,* they can describe false beliefs. Possible worlds can be described in which those propositions could be true, namely, worlds in the mind of the sentence subject. Philosophers call these *propositional attitudes*, in which sentences express false propositions as belonging to another's mind or perspective. While it is true that specific vocabulary words exist to name the state of a false belief, on its own a verb such as "deluded" (3) does not capture the content of the delusion, and the contents of propositional attitudes (4) matter in predicting what the subject will do, or in explaining what they have done.

(3) Miriam's friend was deluded
(4) Miriam's friend thought that Miriam baked the bread.

The properties that sentential complements have for capturing mental states are discussed extensively in the literature on propositional attitudes (Davidson, 1984; Richard, 1990; Segal, 1998).

The special characteristic of complements is their contrast of two perspectives, making them uniquely suited to the representation of propositional attitudes such as belief. There are multiple more subtle differences among complements in English and other languages, representing a network of precise decisions in the learning process for children to master (de Villiers and Roeper, 2016). One difference is whether the embedded clause is tensed: non-finite clauses have neither tense, nor an independent truth-value (5):

(5) Anna wanted him to play in the garden.

The phrase "him to play in the garden" is not something one can evaluate for truth, unlike the final clause in a tensed complement (6):

(6) Anna thought he was playing in the garden.

Children master nonfinite complements (7) before they master tensed, finite complements (8), and do not make "reality" mistakes with the former, even when there is a contrast in what happened (de Villiers et al., 2012),

(7) Anna told him to play in the garden, but he went to sleep instead.

What did Anna tell him to do? – Three-year-olds: play in the garden.

(8) Anna said he was playing in the garden, but he went to sleep instead.

What did Anna say he was doing? – Three-year-olds: sleeping.

De Villiers (1995) argued that much of this grammar is established with verbs of communication, as they share many of the same linguistic distinctions with mental verbs, but the difference is that communication events are overt and not covert as with hidden mental states. The child can understand the reference of a verb like "say," and has undoubtedly heard household arguments about events of speech in which truth can be checked:

(9) You said you fed the dog but you did not!

Think of the mindreading needed instead to produce:

(10) You thought you fed the dog but you did not!

At the end of this acquisition process, the child understands the syntax of finite, non-factive, complements. However, the fine-grained distinctions in meaning among mental verbs-which number in the hundreds-still take time and experience in the world, a process beyond syntax itself, and into discourse pragmatics.

But how could acquiring the structure of complements play a role in establishing the very concept of false belief? It should surely be otherwise, namely that first the child understands false beliefs and second, learns to encode them in complements. Then one would see a correlation between the two achievements, but the lines of causation would be the other way round. Further, many researchers argue in favor of a bidirectional relationship between ToM and language. On the one hand, the child's growing understanding of people's actions, goals, and desires surely provides some conceptual grounding for the meaning involved. Mental states are unobservable entities, so hearing a new word like "think" or "forget" will prompt children to consider what these terms might be labeling, and to conceptualize the distinctions among them (Pyers and Senghas, 2009; Low, 2010). Importantly though, the appearance of these verbs in sentence frames with complements sets limits on their possible meaning (Papafragou, 2001). Could sentential complement structures like (1) have a particular enabling function for false belief reasoning? This particular debate has resulted in much attention, both theoretically (Hinzen, 2007; de Villiers, 2007) and empirically (Boeg Thomsen et al., 2021; Brandt et al., 2023; Farrar et al., 2017; Cheung et al., 2004).

1.3.2 Empirical studies

We know that children begin using verbs such as *think* and *know* from an early age (Bartsch and Wellman, 1995; Diessel and Tomasello, 2001; Shatz, 1994; Shatz et al., 1983) but their first uses may be less like expressions of propositional attitudes than like stereotyped forms, often self-referent, with narrow functions. The forms in (11) and (12) do not capture the contrast between truth and reality, or contrasts across minds, unlike the form in (1).

(11) I do not know (used as an escape from questioning).
(12) I think it's a dog (I think used as "maybe").

Examples like (12) led Diessel and Tomasello (2001) to argue that children treat the high-frequency string "I think" more like an adverb, without explicit mental-state reference. As a result, hearing examples like (12) with first-person subjects might not help children to recognize false-belief understanding. The very first expressions of third person propositional attitudes seem to emerge around 3 or 3.5 years in spontaneous speech, and occur more rarely (Bartsch and Wellman, 1995). However, in experimental settings when children are tested about understanding the forms, consistent difficulty is revealed. For instance, de Villiers (1999) arranged scenarios as in (13) in which characters made statements that were either lies or mistakes, such as:

(13) The woman said she found her slipper. But look, it was really a mouse.

What did the woman say she found?

Three-year-olds consistently answer "mouse," even though the answer is provided in the sentence and one can argue that no "mind reading" is necessary in the situation. Answering correctly entails understanding the discourse and reconstructing what the question refers to, which is only possible if the syntax of complements is mastered. Four-and five-year-olds answer "slipper," as do adults. A longitudinal study of three-and four-year-olds by de Villiers and Pyers (2002) and a very large cross-sectional study of 1,000 children aged four to ten in the standardization of the DELV language assessment test (Seymour et al., 2003) exposed the time course and uniformity of this development (de Villiers et al., 2003).

The finding of correlation between complements and performance on explicit false belief reasoning tasks has been documented now in several different languages: English (e.g., de Villiers and Pyers, 2002; Brandt et al., 2023; Boeg Thomsen et al., 2021) German (Perner et al., 2003; Grosse Wiesmann et al., 2017); Danish (Knuppel et al., 2007), Mandarin (Mo et al., 2014; Guo et al., 2022; Brandt et al., 2023; Li and Leung, 2024), and ASL (Schick et al., 2007). Most but not all of the studies exploring the connection have used the comprehension test originating in the work of de Villiers (1995). In a variant, Brandt et al. (2016) evaluated children's comprehension of complements using first versus third person subjects, and found that only performance with third-person complements correlated with 4-year-olds' false belief performance. Aksu-Koç et al. (2005) found that production of complements in Turkish predicted false belief reasoning better than the production of evidential markers did. In Dutch, de Mulder et al. avoided the truth contrasts of the complement comprehension task and asked the children simply to report what different people reported happening, when the truth of the matter was unknown. Success on that task proved unrelated to false belief reasoning. Yet in Brandt et al., in English, truth contrasts were also side-stepped: the child had to recall what was said in a complement across an intervening sentence that was merely a distractor, but did not deny its truth. In contrast to the results in de Mulder et al., performance on this task was well correlated with false beliefs. In the work in German by Grosse Wiesmann et al. (2017), the child's imitation of complements was used instead, because imitation can betray whether children understand the crucial difference in placement of the complementizer in German. The children's skill at accurate imitation of the complements was connected to their false belief reasoning. Hence, the basic result seems to survive most variations in the task used (but see Brandt et al. (2023) on complement production).

The effects are less clear in children learning Cantonese (Cheung et al., 2004; Cheung et al., 2009; Tardif et al., 2007), a language in which the surface markers of complementation are virtually non-existent and there is no wh-question movement. Tardif et al. (2007) reported a large longitudinal study of children learning Cantonese in Hong Kong, and though she found significant correlations between complement comprehension on the de Villiers and Pyers (2002) "memory for complements" task and false belief understanding, the children overall were surprisingly poor at the complement comprehension test, even at age six. The complements did not seem to be prerequisite for false belief reasoning in Cantonese, a finding echoed in a recent longitudinal path analysis by Siu and Cheung (2022). Yet in closely-related Mandarin, Guo et al. (2022) present robust evidence of a correlation between complements and false belief reasoning in both typical children and children with autism [see also Li and Leung, 2024 (this issue)].

Brandt et al. (2023) explored the CHILDES data on parent–child conversations in Mandarin and found that the parents, unlike their English counterparts, do not show the frequent use of "I think" as an epistemic adverb. In their experimental study, Brandt et al. (2023) found that discriminating the use of the non-factive verb *falsely think* from *know* complements was related to false belief understanding, regardless of the subject being first or third person. English speaking children, however, were affected by the person of the verb in a parallel study. Following a suggestion by Tardif and Wellman (2000) on the relative lack of discourse on mental states from Chinese parents to children, Mo et al. (2014) trained children on complements using verbs of communication. They found the training paid off in children's improvement on false belief reasoning.

When children are language delayed, either by DLD (Developmental Language Disorder), by autism spectrum disorders, or by deafness, the path of development can be stretched out in a way that allows researchers to see the order of mastery (Durrleman et al. 2016). The case of language-delayed deaf children is especially relevant here, as these children do not usually have other associated developmental problems or neurodivergent minds, but may just have restricted access to the primary input. Children who are born to hearing parents and do not have the benefit of exposure to signed languages often struggle to learn oral language, and are delayed several years as a result. Schick et al. (2007) and Tager-Flusberg and Joseph (2005) studied these children's language and false belief development, and report delays in false belief understanding concomitant with their language delay. Both vocabulary and complement comprehension predicted their false belief understanding, even though the false belief tasks were designed to be as minimally verbal as possible. That is, the language required for the task was ruled out as a variable. Since the linguistic skill is needed equally for non-or low-verbal false belief tasks, it suggests that language is needed to develop the false belief reasoning, not just to understand the classic tasks. In an extreme case in Nicaragua, researchers for several decades have studied the invention of a sign language NSL (Nicaraguan Sign Language) by deaf children sent away to special boarding schools. The first generation of such signers had relatively incomplete grammars, without sophisticated grammatical ways to indicate other people's perspectives. The deficiency was then remedied in successive generations of learners. The developmental inversion is that older learners of NSL can be less competent than younger learners. Pyers (2004) and Pyers and Senghas (2009) found that those signers who could express propositional contents under mental state verbs were able to do false belief reasoning (in a nonverbal test), but those signers who did not, failed them even as adults. Incomplete language development hinders theory of mind development even in adult populations.

1.3.3 Explanatory accounts

Despite the empirical data primarily in support of complements as a factor in false belief development, the explanation is still very much debated. Sentential complements represent one example of complex language, possibly the apex of syntactic success in the preschool years. Perhaps general language development (usually indexed by vocabulary) or general syntactic skill are the real predictors of false belief, and complements are just representative of that stage of acquisition (Ruffman, et al., 2002). Call this the *general language* hypothesis. Some have argued that complements reflect the kind of rich discourse that encourages false belief reasoning, that is, the content of the theory about minds is exposed by such conversation (Harris et al., 2005; Nelson, 2005; Tompkins et al., 2018). This can be called the *discourse content* hypothesis. Alternatively, developing complement structures might scaffold the kind of careful reasoning that the scenarios entail, allowing the child to entertain multiple perspectives on the world and successfully choose among them (for a rich discussion of these and other alternatives, see San Juan and Astington, 2017; de Villiers, 2021). This we call the *complements facilitate reasoning* hypothesis.

Empirical results that might allow differentiation of these alternatives are ambivalent, as not all the studies included both general language and complementation measures in the same investigation. For example, Low (2010) found that understanding sentential complements predicted standard false belief tasks in a cross-sectional sample of English-speaking children, once age, nonverbal ability and implicit false belief scores were controlled, but no alternative indices of language were included. Meta-analyses provide some help with this question, but the results are somewhat mixed. In a meta-analysis in Milligan et al. (2007) considered all the language measures used by studies to that date. Various different language measures, including syntax (29%), general language (27%), semantics (23%), and receptive vocabulary (12%) made significant contributions to false belief understanding. However, memory for complements had the strongest effect size in relation to false belief reasoning, with an effect size of 44%, though such studies were less common. Farrar et al. (2017) provide an updated meta-analysis of those studies that contained data on the role of complements and general language as predictors of false belief reasoning. In 10 of the 18 studies (55%) that compared both, the *general language* hypothesis was supported over and above the specific role of complements. These studies used a wide variety of measures to assess "general language ability," including receptive vocabulary and different measures of syntax development. However, six of these ten negative studies were for Cantonese and Korean. Mental state verbs differ in these Asian languages compared to English, in that the distinction between true and false beliefs is carried lexically in the verb (see Tardif et al., 2007). Nevertheless, the majority of these studies tested complements with communication verbs (except for Cheung, 2006; study 2). Thus Farrar et al. argue that even these cross-linguistic studies on Asian languages can be used to evaluate the relative contribution of complementation versus general language.

Since then, more recent work has used a variety of measures to distinguish the contribution of complements versus more general language, usually taking out the contribution of these skills and then testing whether complements explain added variance. Brandt et al.

10.3389/fpsyg.2024.1455941

(2023) and Boeg Thomsen et al. (2021) both contained language measures, and after taking out the contribution of vocabulary and other syntax, found a significant additional contribution of complements measures to false belief understanding in keeping with the *complements facilitate reasoning* hypothesis. An exception mentioned is De Mulder et al. (2019) in a Dutch study, using a complement task that avoided truth contrasts altogether. That task made no contribution to false belief reasoning relative to other language measures such as vocabulary and general syntax. Hence the contribution made by the truth contrasts of non-factive complements is still a live issue (see also de Villiers et al., 2012).

Longitudinal studies are rarer, but they can help identify the direction of influence between the variables, as well as control for the level of initial false belief reasoning. It would be natural to propose that children at three or four years do not yet have the conceptual resources to consider others' perspectives and mental worlds, leading to errors with false complements as a result of their failures to understand others' false beliefs. De Villiers and Pyers (2002) studied a small group ($N = 28$) of children over a year in preschool, and tested them at four points on a battery of theory of mind and language tests. Though they had begun the study expecting that false belief understanding might be necessary for comprehending complementation, the reverse turned out to be the case. When children acquired an understanding of sentential complements, then they began to systematically pass false belief tasks. As a control, de Villiers and Pyers took a language sample and looked at the children's general command of syntax using the IPSYN (Index of Productive Syntax, Scarborough, 1990), an index of the variety and complexity of sentence use. This measure was not as well linked to false belief mastery.

Boeg Thomsen et al. (2021) did a small longitudinal study with 45 English speaking children designed specifically to try to tease apart the potential contribution of two aspects of complement-clause acquisition. As discussed, their complement comprehension task varied from the more common one by not asking complex questions, and by entailing a distractor clause that did not contrast truth. They tested proficiency with the perspective-marking syntactic structure itself, and in a separate task, understanding of the specific mental verbs used, namely *know* versus *think*. As additional language tests, they included memory span, a test of vocabulary, and receptive grammar, and epistemic modals. Furthermore, they avoided complement structures in the false belief task altogether making this one of the most stringent tests to date. They found a robust effect linking third person complements to false belief reasoning when controlling for the rest, as well as an additional effect of mental verb understanding. These studies also are in keeping with the *complements facilitate reasoning* hypothesis.

1.4 Executive function and complements

Given two potential explanations of the changes in theory of mind reasoning in the later preschool years, is it possible to decide between the *complements facilitate reasoning hypothesis* and the *executive function* hypothesis? Some have argued that the complement comprehension task (de Villiers and Pyers, 2002) might in itself be an inhibitory control task, in that to answer the critical question, the reality response must be suppressed (Boeg Thomsen et al., 2021). Unfortunately, it is not often that language and executive function are pitted against one another in the same study. In the cross-sectional study by Jenkins and Astington (1996), there was no longer any effect of working memory (a component of EF) in their 3-and 4-year olds' false belief understanding when controlling for language. Similarly, in Hughes (1998), taking out verbal skills reduced the effects of EF (working memory and inhibitory control) on false belief reasoning in their sample. However, those studies did not consider complements. Low (2010) found that both complement comprehension and executive functioning predicted explicit false belief reasoning in preschoolers, but he could not separate the effects of each of those variables in regression analyses.

Studying children with deafness and delayed language can provide more perspective here. In de Villiers and de Villiers (2012) the oral deaf children were on par with their hearing peers on executive function tasks, but they showed significant delays in false belief tasks even when the language demands of the task were minimized. They report a dissociation of deception and false belief tasks, in that the deaf children were equivalent to their hearing peers on deception games. Furthermore, language, and in particular complement syntax, proved to be the best predictors of false belief reasoning; however, executive function skills, especially inhibitory control, were the best predictors of deception.

A longitudinal study by Farrant et al. (2012) was rich enough to explore the relative contributions of vocabulary, general language, executive functioning and complements to false belief reasoning in English-speaking preschool children. Farrant et al. added to the model the variable of mind-mindedness (Meins et al., 2002), predicting that variation in maternal input about mental states (via perhaps the *discourse content* hypothesis) would predict children's ability on sentence complements, which would then predict false belief understanding. Their sample included 91 typically-developing Australian children studied twice across a year. Importantly, the effects of variation in maternal mental talk were completely mediated by the children's own competence at sentential complements, which predicted their belief ability. Cognitive flexibility was a further predictor, and the direction of effect was that sentential complements predicted this executive function index rather than vice versa. However, the Farrant study had a relatively small sample size for the number of variables, and they did not use structural equation modeling for the longitudinal portion of their study. The recent study by Boeg Thomsen et al. (2021) also tested the effects of inhibitory control as well as several other language variables in their longitudinal work, but there is a limit on the power of linear regressions with so few participants (45) and so many test variables. The regressions revealed a bidirectional effect of complements and false belief reasoning across time.

In sum, the existing studies are typically too small and contain too many variables for a sufficiently powerful statistical technique such as structural equation modeling that can differentiate direction of effects in a longitudinal study.

1.5 The current study

In the current study we had the opportunity to test the different theoretical models of false belief development on a large sample of low-income children studied over the course of several years on a research grant attached to a preschool curricular intervention study (Lonigan et al., 2015). In our project the children received a large battery of language, executive function, and theory of mind measures, and these were repeated several times over the course of the study, making this an ideal group to test competing models.

Lonigan et al. (2015) is a report on the curriculum intervention, using some measures that overlapped with the current study but it does not include the measures of complex language, executive function or theory of mind used here.

The intervention study itself was a large-scale cluster-randomized investigation of the effects of an integrated literacy-and math-focused preschool curriculum. The basic curriculum incorporated central elements of the Literacy Express Curriculum (Lonigan et al., 2005) and Pre-K Mathematics (Klein et al., 2002), two preschool curricula rated as effective by the US Department of Education's What Works Clearinghouse. One hundred and ten center-based preschools serving low-income communities in the Houston, TX and Tallahassee, FL areas were randomly assigned to one of three conditions: 1. the base curriculum with added explicit socio-emotional instruction [the Promoting Alternative Strategies Thinking curriculum (PATHS) (Domitrovich et al., 2007)]; 2. the base literacy and math curriculum and general classroom and behavior management instruction for the teachers, but with no explicit lessons and activities targeting socio-emotional skills (called the *Implicit* Socio-Emotional Condition); or 3. a "business-as-usual" control condition. There were significant positive impacts of the two versions of the curriculum on language, phonological awareness, math, and socioemotional outcomes. There were no added benefits to academic or socioemotional outcomes for the children receiving *explicit* socioemotional instruction (Lonigan et al., 2015). For this reason, in the current research we combine the first two conditions into the Intervention Condition and compare that to the Control Condition to explore the effects of the intervention on false belief reasoning in our sample of children.

Our study is unique in using a longitudinal design with a very large N (258 preschoolers), allowing for powerful statistical analyses to separate the effects of different possible predictors of the development of false belief reasoning. Both the familiar verbal false belief reasoning tasks (unseen object displacement and unexpected contents) and low verbal tasks (thought-bubble picture narratives) were employed, with enough false belief questions (14) to provide considerable variance in the dependent variable to tease apart effects of different independent variables. Most importantly, longitudinal measures of inhibitory control and several relevant measures of language acquisition were taken. The language measures included expressive vocabulary, morphosyntax, and complement comprehension and there were enough items in each measure to provide sufficient variance to separate the effects, if any, of each aspect of language. Prior studies of the effects of children's language acquisition on their false belief reasoning have often had too few subjects for reliable statistical analyses, have taken too few measures of false belief reasoning, have not pitted executive functioning against language, or have not separated the effects of different aspects of language.

1.6 Predictions

Based on our review of past work, we make the following predictions:

(1) Curricular intervention that enriches socio-emotional understanding is expected to improve false belief reasoning.
(2) Inhibitory control will predict false belief reasoning concurrently but not across time when other variables such as language are controlled.
(3) Language measures will predict false belief understanding both concurrently and over time.
(4) Complement comprehension will be a significant predictor of false belief reasoning over and above other control variables, including more general language measures
(5) Complement comprehension will be correlated with inhibitory control, but complements will be a more important proximal predictor of false belief reasoning compared to inhibitory control.

TABLE 1 Sample by ethnicity.

Ethnicity	N (%)
African American	141 (54.7)
White	78 (30.2)
Hispanic	26 (10.1)
Other	13 (5.0)

TABLE 2 Gender make-up of sample.

Gender	N (%)
Male	126 (48.8)
Female	132 (51.2)

2 Method

2.1 Participants

The participants were 258 children from subsidized[1] preschools in Texas and Florida. 90.4% of the children came from low-income families and were eligible for free lunches. Tables 1 and 2 give the sample statistics on gender and ethnicity. They were recruited as a part of a curriculum intervention project funded by the National Institutes of Health, the School Readiness Research Consortium. All of the children were assessed twice on a battery of socio-emotional, cognitive, quantitative, language and pre-reading tasks during the preschool year. The first time was in September at the beginning of the school year before the curriculum intervention (Time 1) and the second time was in April/May at the end of the school year, after the intervention (Time 3). The children ranged from 3.3 to 5.3 years old (Mean = 4.58, sd. = 0.34) at Time 1 and from 3.8 to 5.9 years old (Mean = 5.11, sd. = 0.34) at Time 3.

Out of 760 preschoolers who completed Time 3 of the NIH curriculum project, the 258 children in the present study included only those children who completed all of the verbal and low verbal false belief tasks as well as all of the relevant tests of inhibitory control and language acquisition at both Time 1 and Time 3. We thus avoided the issue of imputation of missing data. Given the intensity of the testing schedule across four projects, some loss of data was inevitable,

1 In the US, Early Education and Care (EEC) provides subsidies to families in need to help cover the cost of child care. Child care programs accepting these subsidies use EEC's Child Care Financial Assistance (CCFA) system for documenting family eligibility, tracking child attendance, and submitting requests for reimbursement.

especially as these were low-income day care centers with considerable changeover and parents who moved a lot. However, there is no evidence that the children that we excluded for missing tests (170) constituted a different population than those who completed all the measures (258). We checked this by comparing the two groups on the vocabulary measure, which was present for 100% of the sample, and also age. The groups did not differ statistically on either measure. Children who were not monolingual speakers of English were also excluded. Parents had to return a questionnaire on home language use to determine if the child was monolingual in English (183 of these were missing). Children who were reported by their primary caregiver as having more than 10% Spanish input and/or use were excluded from the present study to eliminate possible effects of bilingualism. There were 149 such exclusions.

2.2 Procedure

The children were tested individually on the full assessment battery in three one-hour testing sessions. The testing took place over several days, contingent on the availability of particular children in the center on a given day, and because the overall test battery including this study was very large. Tests were given in a fixed order across the children in each of these sessions.

2.2.1 Measures for the present study

Background measures

Nonverbal IQ

Children's nonverbal IQ was assessed using the Pattern Analysis subtest of the Stanford-Binet IQ test, which measured children's capacity to recognize abstract visual patterns and to solve pattern matching problems. This test was carried out at Time 2, midway between Time 1 and Time 3.

Verbal memory

Verbal memory was assessed with the Word Span subtest on the Comprehensive Test of Phonological Processing (CTOPP) (Wagner et al., 1999). The children were asked to repeat a string of familiar English words said quickly by the tester (e.g., "fish, bed, dog") in the same order as they had been produced. The strings of words increased in length until the child failed to repeat them accurately. This test was also administered at Time 2.

False belief reasoning

Verbal FB reasoning tasks

Unexpected contents

The unexpected contents task is a standard task about whether children can remember their earlier false belief about the expected contents of a familiar container after knowing the true content, and whether they can predict others' false belief about the container's contents (Perner et al., 1987). Each participant was shown a box (e.g., a crayon box) and the tester asked them what they believed to be in the box. After the child's answer, the tester opened the box and showed that it actually contained an unexpected item (e.g., a spoon). The child was then asked what they thought was inside the box before it was opened, and what their friend or a toy character would think before it was opened. Each child was tested on two different containers at T1 and another two containers at T3, so they could not just remember what was in the container from T1 to T3. Our prior research over some twenty years had shown the four containers with unexpected contents to produce essentially equivalent responses from preschoolers.

Unseen object displacement

The unseen object displacement task is a modification of the Sally-Anne test (Wimmer and Perner, 1983), which is the standard task measuring whether the children can predict a character's behavior after successfully reasoning about their false belief (de Villiers and de Villiers, 2012). Children were told two different picture-supported stories, each involving two characters. One character (e.g., a boy) put an object (e.g., a basketball) at one place and left the room. The other character, without deceptive intent, put the object somewhere else. The children were then asked where the boy would first look for the basketball when he came back into the room and why he would look there for the ball. Different stories were used at T1 and T3. Prior research had shown that the four stories produced very similar responses from preschoolers.

Children's performance received a score of either 1 (for a correct response) or 0 (for an incorrect response) for each verbal false belief question across the Unexpected Contents and Unseen Displacement tasks. This produced a total verbal false belief score out of 8 for each participant.

Low-verbal FB tasks

Thought bubbles tasks

Both 2-picture sequences (Woolfe et al., 2002) and 4-picture sequences (de Villiers and de Villiers, 2012) were employed. The low-verbal tasks were tested together. Three training items were completed before the administration of low-verbal tasks to make sure the children understood the nature of thought-bubbles as representing the contents of what a character was thinking. Children who failed the training would not be given the low-verbal tasks.

The 2-picture sequence procedure (Woolfe et al., 2002) contained three false belief trials and two true belief trial. The children were shown pictures in which a character was thinking something but the thought bubble was empty, and were asked to choose from three pictured objects what they thought should go into the bubble. An example of a false belief trial showed a picture of a boy fishing with a bending rod. The end of the line and the hook were covered by a flap. The tester instructed the child to lift the flap and a boot was revealed to the child but not the character in the picture. Then the tester pointed to the thought bubble on the boy's head and asked "What goes it here?." The child would then choose among three pictures: a fish (false belief), a boot (reality) and a bird (distraction). An example of a true belief trial would be identical to the false belief trials except it would reveal a fish rather than a boot.

The 4-picture-sequence procedure (de Villiers and de Villiers, 2012) included three false belief trials and one true belief trial. In these picture sequences, one of the two characters replaced the contents of a familiar container (e.g., playdough) with an unusual object (e.g., worms). A second character either observed or did not observe this replacement. The children were then asked what should be put into the thought bubble above the second character's head before s/he opened the container (e.g., a picture of playdough or a picture of worms). The picture was drawn in a way such that it was reasonable for the character to want to find either object that could be in the container.

For each false belief trial, the child received a score of 1 (for a correct answer) or 0 (for an incorrect answer). The total score out of 6 false belief trials was then recorded as the child's low-verbal false belief score.

Language

Vocabulary

Children's achieved vocabulary in English was assessed with the Expressive One Word Picture Vocabulary Test (EOWPVT-R; Brownell, 2000) at both T1 and T3. They were asked to name the objects, actions, or concepts shown in colored pictures. In all our analyses we used the children's raw scores on the test, the number of words correctly produced.

Morphosyntax

As an additional measure of the children's general language development, we used their raw scores on the morphosyntax items of the Risk Subtest of the Diagnostic Evaluation of Language Variation – Screening Test (DELV-ST; Seymour et al., 2003). This test was used because it assesses African-American children's acquisition of morphosyntax that is neutral with respect to differences between African American English (AAE) and Mainstream American English (MAE). We did not wish to use standardized tests of MAE that are biased against children speaking other dialects of English, since a high proportion of the children in our study were African American (See Table 1), many of them AAE speakers.

False complement comprehension

Children's ability to process syntactic structures that differentiate what a person said from reality was tested using the task developed by de Villiers and Pyers (2002). The False Complement Comprehension task contains 8 trials testing whether the children can hold the complement structure associated with the verb "say" in memory and successfully reproduce it. In each trial, a story was told along with pictures about one character saying something which was different from reality. For example, "The woman said there was a bug in her cereal. But look, it was just a raisin!" At the end of the story the tester asked "What did the woman say was in her cereal?" For each trial, the child received a score of either 1 (passed) or 0 (failed). The total score out of 8 was then recorded.

Inhibitory control

Bird and dragon task

The Bird and Dragon Task is a simplified version of "Simon Says" (Reed et al., 1984; Kochanska et al., 1996). In this task children need to selectively inhibit commanded actions. In 10 familiarization trials, the tester first asked the children to imitate self-directed action (e.g., "Touch your ears"). Then the tester introduced two puppets: a "nice bird" and a "naughty dragon." The instruction was to follow the bird's commands but not the dragon's. Children's performance on each dragon command received a score of 1 (performed a full movement), 2 (performed a wrong movement), 3 (performed a partial movement), or 4 (did not move), and the mean of scores on 7 dragon commands was recorded as the "inhibition score."

Knock-tap task

The Knock-Tap Task required children to be able to switch from imitating hand actions to doing the opposite action (Korkman et al., 1998). First the children were asked to imitate the examiner by either knocking with a closed fist or tapping with an open palm on a box for 8 trials. Then for 8 pseudorandom opposite trials the children had to tap when the examiner knocked and knock when the examiner tapped. Thus, in this task the children had to inhibit the prepotent response of imitating the tester's hand action, the response that had just been primed. Percentage of correct responses over 8 opposite trials was recorded for each child.

3 Results

3.1 Background measures

Our analysis of the factors contributing to false belief reasoning in the children began with two regression analyses. The first looks at the cross-sectional predictors at T3, taking out the level of FB at T1. The second takes advantage of the longitudinal data to look also at predictors from T1 to T3. Finally, we construct a structural equation model to look at the interactive effect of the predictors. But first an outcome score, a composite of false belief measures was required. Children's scores on the four false belief tasks were significantly intercorrelated (see Table 3), so a composite Total FB score was created by adding together the responses to all of the FB questions. Scores varied from 0 to 14 correct out of 14.

Although we intended to create a composite executive function score, a composite inhibitory control score was not calculated from the Bear-Dragon and Knock-Tap Test because the partial correlations between performances on the two measures were low ($r = 0.133$, $p = 0.033$ at T1; $r = 0.114$, $p = 0.069$ at T3; partial correlations

TABLE 3 Intercorrelations between the different False Belief tasks at Time 3 ($N = 258$).

	Unseen Displacement	Unexpected Contents	Thought Bubble 2-Picture
Unexpected contents	0.316***		
Thought bubble 2-Picture	0.307***	0.296***	
Thought bubble 4-Picture	0.360***	0.312***	0.380***

TABLE 4 Longitudinal growth in false belief reasoning, inhibitory control, and language measures between Time 1 (early in the preschool year) and Time 3 (late in the preschool year).

Measure	Time 1	Time 3	*t*	df	*p*
ToM					
False Belief	5.90 (3.24)	8.98 (3.02)	17.50	258	0.000***
Inhibitory control					
Bear-Dragon	0.74 (0.35)	0.93 (0.18)	8.49	258	0.001**
Knock-Tap	0.68 (0.34)	0.82 (0.26)	5.64	258	0.002**
Language					
Morphosyntax (DELV-ST)	5.63 (1.12)	6.27 (0.085)	11.07	258	0.000***
Vocabulary (EOWPVT-R)	38.75 (11.23)	46.60 (12.09)	15.11	258	0.000***
Complements	6.02 (2.05)	7.34 (1.34)	11.63	258	0.000***

TABLE 5 Cross-sectional hierarchical linear regression predicting T3 False Belief from Background variables (age, gender and intervention), T1 false belief scores, inhibitory control, and language measures.

	ΔR^2	*F*(df)	*p*	Predictor	ß	*t*	*p*
Background	0.157	15.80 (3,254)	0.000***	T3 Age	0.29	5.03	0.000***
				Gender	0.188	3.27	0.001**
				Intervention	0.184	3.2	0.002**
T1 ToM	0.247	42.96 (1,253)	0.000***	T1 False Belief	0.536	10.25	0.000***
Inhibitory control	0.009	3.85 (1,252)	0.050*	T3 Bear-Dragon	0.097	1.96	0.050*
Language	0.127	22.94 (3,249)	0.000***	T3 Morphosyntax	−0.049	−0.92	ns
				T3 Vocabulary	0.204	3.86	0.000***
				T3 Complements	0.306	6.06	0.000***

controlling for Age). The Bear-Dragon test was selected as the predictor because it had the strongest bivariate correlation with False Belief.

The children showed significant longitudinal growth over the 6 months between T1 and T3 in their false belief reasoning, in each of the Inhibitory Control tasks, and in the three Language measures (see Table 4).

There were significant effects of both Gender and Curriculum Intervention on the children's total FB score at T3 (Gender: F (N = 132) FB mean = 9.53/14, M (N = 126) FB mean = 8.40/14; $t(256) = 3.03$, $p = 0.003$) (Intervention: INT (N = 215) FB mean = 9.25/14, CONTROL (N = 43) FB mean = 7.63/14; $t(256) = 3.27$, $p = 0.001$). Therefore, both Gender and Intervention were entered as background predictor variables in the subsequent analyses of the relationship between inhibitory control and language development and the children's FB reasoning toward the end of preschool. Predictors of Total FB reasoning scores at T3 were examined using both linear hierarchical multiple regressions and structural equation modeling.

3.2 Cross-sectional predictors of false belief reasoning at time 3: regression 1

First, a cross-sectional multiple regression was performed relating T3 raw scores on the Bear-Dragon Game and the three Language measures to the children's T3 Total FB scores. The Language measures were kept separate to investigate the independent effects of Morphosyntax, Vocabulary, and Complements on FB reasoning, although the regression also showed the combined effects of the three on the variance in total FB score at T3. Predictors were entered in four blocks: first, Background Variables [Age at T3, Gender (1 = Female, 0 = Male), and Intervention (1 = Intervention Group, 0 = Control Group)]; second, the children's T1 FB scores; third, performance on the Bear-Dragon task; and finally, the three Language measures (see Table 5).

The Background variables accounted for a significant proportion of the variance in T3 FB scores ($R^2 = 0.157$, $p = 0.000$) and each of the three variables was a significant contributor (see Table 5). Children's FB scores at T1 were major predictors of their FB reasoning at T3, accounting for an additional 24.7% of the variance ($p < 0.001$). Performance on the Bear-Dragon Game at T3 was also a significant predictor of later FB reasoning ($\Delta R^2 = 0.009$, $p = 0.050$), although the additional variance accounted for by this measure was small. Finally, the T3 Language measures added substantially to the prediction of T3 Total FB scores ($\Delta R^2 = 0.127$, $p = 0.000$). Of the Language measures, both Expressive Vocabulary and Complement Comprehension were significant independent predictors of T3 FB (Vocabulary: $B = 0.204$, $t = 3.86$, $p < 0.001$; Complements: $B = 0.306$, $t = 6.60$, $p < 0.001$), but Morphosyntax was not in itself a significant predictor.

TABLE 6 Longitudinal hierarchical linear regression predicting T3 false belief from background variables (age, gender and intervention), T1 false belief scores, inhibitory control, and language measures.

	ΔR^2	*F*(df)	*p*	Predictor	ß	*t*	*p*
Background	0.157	15.80 (3,254)	0.000***	T3 Age	0.29	5.03	0.000***
				Gender	0.188	3.27	0.001**
				Intervention	0.184	3.2	0.002**
T1 FB	0.247	42.96 (1,253)	0.000***	T1 False Belief	0.536	10.25	0.000***
Inhibitory Control	0.03	13.39 (1,252)	0.000***	T1 Bear-Dragon	0.188	3.66	0.000***
Language	0.071	11.93 (3,249)	0.000***	T1 Morphosyntax	0.004	0.07	ns
				T1 Vocabulary	0.261	4.65	0.000***
				T1 Complements	0.138	2.55	0.011*

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

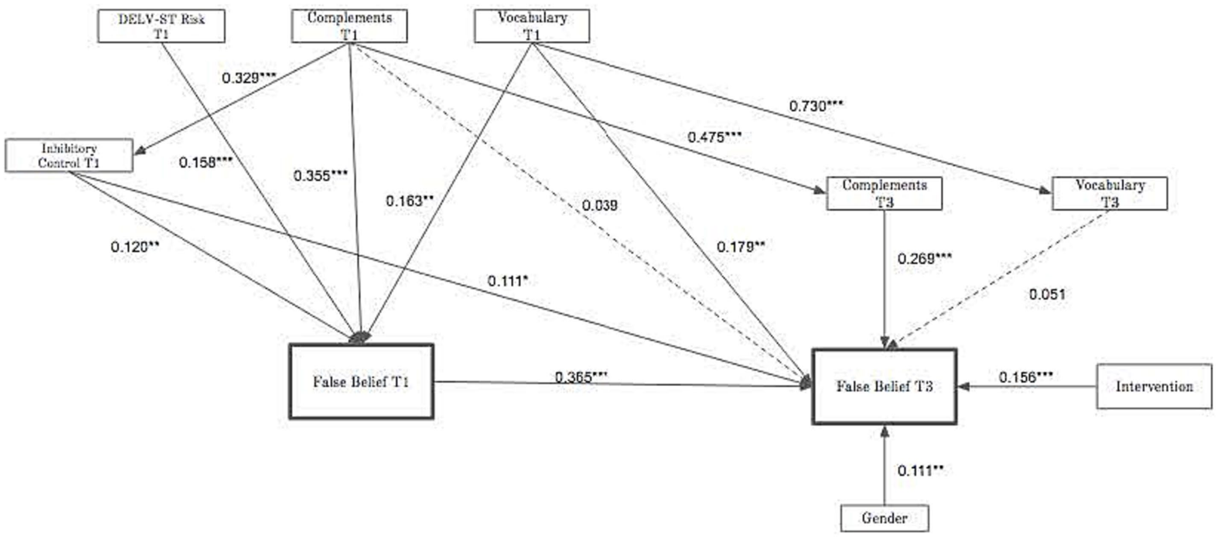


FIGURE 1
Concurrent and longitudinal structural equation model of the relationships between inhibitory control (Bear-Dragon game), the three language measures, and total false belief scores at Tl and T3. Effects of gender and curriculum intervention are also part of the model. Raw scores were entered for the Bear-Dragon game, DELV-ST Morphosyntax, complement comprehension, and vocabulary at Tl and T3 . The figure gives standardized parameter estimates for the hypothesized longitudinal SEM. All reported estimates are the maximum likelihood standardized point estimates. χ^2 (21, *N* = 258) = 38.008, $p = 0.013$; Comparative Fit Index = 0.978; Tucker-Lewis Index = 0.954; Root Mean Square Error of Approximation (RMSEA) = 0.056; Standardized Root Mean Square Residual (SRMSR) = 0.068. All error covariances were calculated but are not shown.

3.3 Longitudinal predictors of false belief reasoning at time 3: regression 2

The second regression considered the possible longitudinal effects of the additional measures taken at T1, not just the continuity in FB scores from T1 to T3. For the longitudinal regression analysis between T1 and T3, the predictor variables were also entered in four blocks (Table 6). First, the three background variables were entered: Age at T3, Gender, and Intervention. Next, the children's T1 FB scores were entered. Then scores on the Bear-Dragon Inhibitory Control game at T1 were entered. Finally, the three Language measures from T1 were entered as a block: DELV-ST dialect neutral Morphosyntax, Expressive Vocabulary, and Complement Comprehension. Thus, the first two blocks of this regression analysis were the same as the cross-sectional T3 analysis, but longitudinal Inhibitory Control and Language predictors were entered into this regression.

The results for Block 3 and Block 4 of the longitudinal regression were very similar to those from the cross-sectional analysis: performance on the T1 Bear-Dragon game added significantly to the variance in T3 FB scores accounted for by the variables in Blocks 1 and 2 ($\Delta R^2 = 0.030$, $p = 0.000$) and Expressive Vocabulary and Complement Comprehension at T1 were each independent predictors of later T3 FB reasoning (Vocabulary: B = 0.261, $t = 4.65, p < 0.001$; Complements: B = 0.138, $t = 2.55$, $p < 0.011$). The Language measures as a group accounted for 7.1% of the variance in FB scores ($p < 0.0001$). Morphosyntax scores on the DELV-ST did not add any predictiveness to the Language measures.

3.4 Structural equation model

The power of structural equation methods lies in their ability to model complex interactions and direction of effects, especially in longitudinal data. A structural equation model was fitted to the data from T1 and T3 using the same predictor variables and outcome measures as the regressions (see Figure 1). The model shows both concurrent effects of Inhibitory Control and Language measures at T1 and T3 on FB scores at T1 and T3, and longitudinal effects of the T1 measures on T3 scores. The final model is an excellent fit to the data, with fit indices >0.95 and RMSEA and SRMSR values <0.07.

The SEM complements the findings of the two regression analyses. Inhibitory Control (Bear-Dragon) at T1 has significant direct effects on FB at T1 (Standardized Parameter (SP) = 0.12, $p < 0.001$) and longitudinally at T3 (SP = 0.111, $p < 0.05$). Bear-Dragon scores at T3 did not add to the fit of the model and standardized parameter estimates between that measure and FB at T3 were not statistically significant, so that variable was excluded from the final model.

Of the Language measures, Vocabulary and Complement Comprehension have stronger effects than Morphosyntax. Morphosyntax scores on the DELV-ST at T1 are significantly related to FB at T1 (SP = 0.158, $p < 0.001$), but not to FB at T3. T3 Morphosyntax scores did not have significant effects on T3 FB, so they are excluded from the model.

Complement Comprehension scores have direct effects on FB at T1 (SP = 0.355, $p < 0.001$), but not longitudinally on FB at T3. Instead, they have strong indirect effects on T3 FB through their effects on T3 Complement Comprehension (SP = 0.475, $p < 0.001$). T3 Complement Comprehension is significantly related to FB at T3 (SP = 0.269, $p < 0.001$).

Expressive Vocabulary at T1 has significant direct effects on both T1 (SP = 0.163, $p < 0.001$) and T3 FB (SP = 0.179, $p < 0.001$), and T1 Vocabulary has a significant longitudinal effect on T3 Vocabulary (SP = 0.730, $p < 0.001$); but T3 Vocabulary is not significantly related to T3 FB in this model. The best fit directional pathway between Inhibitory Control at T1 and Complement Comprehension at T1 goes from Complements to Inhibitory Control (SP = 0.329, $p < 0.001$).

4 Discussion

The data analyses allow separation of the various influences on false belief reasoning over time, taking into account the child's initial level of success. Of course, these results do not rule out the possibility that different developmental paths might occur in other cultures and other languages, which is why a broader net of research must still be cast. This is the first study of a large sample of poorly-resourced children in the US, the majority of whom were African American or Hispanic.

Our first prediction concerned the effects of the intervention on false belief reasoning. Since the interventions were randomized, all the groups contained an equivalent mix of children from low-income households with parents of low educational achievement. The larger group of researchers had an interest in curricular improvements that could be implemented in low-income day centers that could pay off in early schooling, and as such, were not focused on either syntax nor theory of mind (Lonigan et al., 2015). However, the intensive exposure to books, exercises and materials that focused on emotion regulation undoubtedly enriched the children's exposure to language about the mind, and it paid off in the strong effects of the intervention versus business-as-usual uncovered in the analyses. These results cannot distinguish between the various hypotheses, but are at least compatible with the *discourse content* hypothesis, namely, that the children's development of theory of mind is enriched by more mind-talk (Cutting and Dunn, 1999; Meins et al., 2002). They reinforce and refine the Lonigan et al. (2015) results in that they demonstrate an effect of teacher-delivered, richer day care curricula on theory of mind development in the least-resourced children.

The finding that the intervention impacted theory of mind outcomes for the children was not previously reported, as the published report from the longitudinal study looked only at emotion understanding and parental reports of social development (Lonigan et al., 2015). Including all the groups may have increased the variance in achievements of the sample. The regression already accounted for the variance due to intervention, age and gender, and the language variables then entered were additional sources of influence.

Like some other studies, we also found an effect of gender on false belief reasoning, a fact that is not fully explained. Some prior studies with preschool-aged children have found girls to show slightly better performance on emotion understanding and false belief tasks (Banerjee, 1997; Charman and Clements, 2002), and some found differences in later theory of mind skills (Calero et al., 2013; Kuhnert et al., 2017). However, several others have found no difference (Devine and Hughes, 2012; Mathieson and Banerjee, 2011; Walker, 2005). A gender effect has sometimes been attributed to the differential discourse of parents by the gender of their child, with girls receiving a higher volume of talk about people and relationships, hence mental states (Charman and Clements, 2002). The other possibility is that it reflects the finding that girls outpace boys in language development, though that difference has usually been found to be limited to the early years. In the current data, the girls showed small but statistically significant differences in each of the language skills relative to the boys.

Turning to the theoretical significance of the findings, one can see reflections here of all the major theoretical proposals of factors that influence development. The results of the SEM turned out to be more subtle than our initial predictions. First, it is evident that the child's *executive functioning*, indexed here by inhibitory control, exerts differential influence across ages. Whether measured at T1 or at T3, inhibitory control has a small but significant effect on false belief reasoning at T3 as predicted. However, the SEM model makes clear that the variance in inhibitory control at T1 has its significant influence by affecting false beliefs at T1. By T3, inhibitory control makes no difference, primarily because the variance at T3 is so limited: the majority get full points on the measure. Recall also that our measure was limited to a single task, because of the lack of correlation between the two inhibitory control tasks with which we began. Surprisingly, inhibitory control was not a contributor to complement comprehension: the SEM makes clear that the direction of influence is the reverse. The complement comprehension test has been seen as potentially a kind of inhibitory control task, but that is not evident here. Farrant et al. (2012) found the same direction of effect in their SEM model with typical children, namely, that sentential complements predicted executive function rather than vice versa.

A picture is beginning to emerge of how the different factors contribute to the two processes involved in explicit false belief reasoning. Inhibitory control would seem to be needed for the second step, of suppressing the prepotent response. If the representation of alternatives is appropriate then inhibitory control can assist, but if the representation is inadequate then inhibitory control will not help. The representation must be built first, and language, specifically complementation, is the ideal structure. A child could have developed inhibitory control for other decisions but still fail on both complements and on false beliefs, as in the case of the oral deaf children in de Villiers and de Villiers (2012). A child could theoretically have complements but not inhibitory control, perhaps a child with ADHD, though this has not been explored. The interpretation is partly supported by the results of a recent training study (Boeg Thomsen et al., 2024) that demonstrates that preschool children benefited most from complement training when they had better executive function skills to begin with.

Second, there is support here for the *general language* hypothesis such as vocabulary development or general morphosyntax, as predicted. The child's vocabulary at T1 matters for false belief understanding both at T1 and at T3, but like inhibitory control, the effect of T3 vocabulary on false belief understanding at T3 disappears, and a similar pattern occurs with morphosyntax, despite there being plenty of variance in both still at T3. A major effect at both time points is complement comprehension. What matters at T3 is complement comprehension at T3. That is, the proximal effect by the final stage is the children's complement comprehension measured then. The other influences at an earlier point affect the earlier stages of false belief understanding comprehension. The average false belief understanding score at T1 was 5.46, well below chance for 14 items ($p = 0.001$), but by T3, it was 10.52 (above chance, $p = 0.001$) with a distribution that shows more equal proportions of passers and failers. It is therefore possible that vocabulary, morphosyntax and inhibitory control each make a difference in getting any points at all on the false belief task, but once that is underway, it is complement comprehension that matters. Of course, the present study is still limited by the range of language measures used, and it remains possible that other complex language could play a role in enhancing false belief reasoning (Schroeder et al., 2021; Charnavel et al., 2024).

The importance of these findings is that they reveal how different studies, with small and slightly different age samples, can come up with different results in which executive function, vocabulary, or complement syntax might make different contributions statistically. Furthermore, enriched talk about the mind via specially prepared curricula or books may also be effective in promoting growth. Only with longitudinal data of this degree of richness can a full developmental story emerge. However, though the age span covered here is quite broad (3.3 to 5.9), the longitudinal time interval was still relatively short as the children averaged 4.6 at T1 and 5.1 at T3.

The result that complements predict false belief reasoning confirms the predictions and is compatible with the data from a number of training studies that have set out to test whether false complements can be trained, and if so, whether that linguistic training impacts false belief reasoning. An excellent review of existing training studies is provided in Boeg Thomsen et al. (2024). Though some studies do not conclusively demonstrate a specific effect of complementation (Hale and Tager-Flusberg, 2003; Lohmann and Tomasello, 2003), others provide convincing evidence that specific training on complements has positive effects on later false belief reasoning (Mo et al., 2014; Durrleman et al., 2019; Durrleman and de Villiers, 2024; Boeg Thomsen et al., 2024). Training studies, since they constitute intervention in the causal chain, can be considered the ultimate experimental test of the causal hypothesis, unlike even the best of longitudinal studies that do not train. But training studies do not settle the question of whether this is the normal path of development. Together with reliable longitudinal data, the case is more robust.

Several theorists have maintained that language may be prerequisite for, or an important determinant of, explicit but not implicit theory of mind (Apperly and Butterfill, 2009; Low, 2010). Low and Watts (2013) suggested a signature limit (Apperly and Butterfill, 2009) on the implicit theory of mind system. The proposal is that the implicit system may be capable of tracking only a subset of mental states, such as a protagonist's beliefs about the location of an object, but not beliefs about the object's properties or identity. Results of a study by Fizke et al. (2017) support this proposal. Perhaps language is more influential, or necessary, for beliefs about object identity. In unseen displacement the child must anticipate where a person will look, and that could be done by a system that recognizes another's goals or intentions, but not necessarily beliefs. However, object identity tasks require representing how the other person conceives of an object, and that is a more demanding representation (Rakoczy, 2017). More precise work is needed to test this distinction.

The overall findings of the study are compatible with the claim that sentential complements play a significant role in explicit false belief reasoning development, separate from general morphosyntax, vocabulary and executive function. Many questions remain about the scope of the finding cross linguistically, as so few languages have so far been investigated and the range of mental state expressions is diverse, as are the techniques for assessing them. Questions also remain about the role of implicit theory of mind as an additional component of influence on the variance in children's achievement. More research is needed on whether there is real continuity of skills across the whole span from infancy to higher order belief at age six or seven, and what effects there might be of the type of belief engaged in the task.

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 Smith College Institutional Review Board. 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

JV: Conceptualization, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing. PV: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, 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. The research was supported by a Program grant from the National Institutes of Child Health and Development (NICHD) (PO1 HD048497). The authors received funding from NIH for conducting the research.

Acknowledgments

We gratefully acknowledge the work on structural equation modeling carried out by two students of ours at the time they were at Smith College: Meng Chen, now in a doctoral program at Penn. State University, and Rachel Yan, now in a doctoral program at Michigan University. We also must thank the hard work of everyone in the School Readiness Research Consortium who collected the data: C. J. Lonigan, B. M. Phillips, and J. Clancy-Menchetti (Florida State University); S. Landry, P. Swank, M. Assel, and H. Taylor (University of Texas in Houston).

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

Aksu-Koç, A., Aydın, Ç., Avcı, G., Sefer, N., and Yaşa, Y. (2005). The relation between mental verbs and theory of mind performance: Evidence from Turkish children. Berlin: In IASCL Convention.

Apperly, I. A., and Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychol. Rev.* 116, 953–970. doi: 10.1037/a0016923

Astington, J. A., and Baird, J. (2005). Why language matters for theory of mind. Oxford: Oxford University Press.

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

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

Baldwin, D. (1995). "Understanding the link between joint attention and language" in Joint attention: Its origins and role in development. eds. C. Moore and P. J. Dunham (Hove: Psychology Press).

Banerjee, M. (1997). Hidden emotions: pre-schoolers' knowledge of appearance-reality and emotion display rules. *Soc. Cogn.* 15, 107–132. doi: 10.1521/soco.1997.15.2.107

Bartsch, K., and Wellman, H. M. (1995). Children talk about the mind. Oxford: Oxford University Press.

Boeg Thomsen, D., Kandemirci, B., Theakston, A., and Brandt, S. (2024). Do complement clauses with first-or third-person perspective support false-belief reasoning? A training study with English-speaking 3-year-olds. *Dev. Psychol.* doi: 10.1037/dev0001808

Boeg Thomsen, D., Theakston, A., Kandemirci, B., and Brandt, S. (2021). Do complement clauses really support false-belief resoning? A longitudinal study with English-speaking 2-to 3-year-olds. *Dev. Psychol.* 57, 1210–1227. doi: 10.1037/dev0001012

Brandt, S., Buttelmann, D., Lieven, E., and Tomasello, M. (2016). Children's understanding of first and third person perspectives in complement clauses and false belief tasks. *J. Exp. Child Psychol.* 151, 131–143. doi: 10.1016/j.jecp.2016.03.004

Brandt, S., Hargreaves, S., and Theakston, A. (2023). Putting complement clauses into context: testing the effects of story context, false-belief understanding, and syntactic form on children's and adults' comprehension and production of complement clauses. *Cogn. Sci.* 47:e13311. doi: 10.1111/cogs.13311

Brandt, S., Li, H., and Chan, A. (2023). What makes a complement false? Looking at the effects of verbal semantics and perspective in mandarin children's interpretation of complement-clause constructions and their false-belief understanding. *Cogn. Ling.* 99–132. doi: 10.1515/cog-2021-0108

Brooks, R., and Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: a longitudinal study connecting gaze following, language, and explicit theory of mind. *J. Exp. Child Psychol.* 130, 67–78. doi: 10.1016/j.jecp.2014.09.010

Brownell, R. (2000). Expressive one-word picture vocabulary test. *3rd* Edn. Novato, CA: Academic Therapy.

Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112, 337–342. doi: 10.1016/j.cognition.2009.05.006

Calero, C. I., Salles, A., Semelman, M., and Sigman, M. (2013). Age and gender dependent development of theory of mind in 6-to 8-years old children. *Front. Hum. Neurosci.* 7:281. doi: 10.3389/fnhum.2013.00281

Call, J., and Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci.* 12, 187–192. doi: 10.1016/j.tics.2008.02.010

Carlson, S. M., Claxton, L. J., and Moses, L. J. (2015). The relation between executive function and theory of mind is more than skin deep. *J. Cogn. Dev.* 16, 186–197. doi: 10.1080/15248372.2013.824883

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

Carlson, S. M., Moses, L. J., and Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant Child Dev.* 11, 73–92. doi: 10.1002/icd.298

Charman, T., and Clements, W. (2002). Is there a gender difference in false belief development? *Soc. Dev.* 11, 1–10. doi: 10.1111/1467-9507.00183

Charnavel, I., de Villiers, J. G., and Durrleman, S. (2024) Taking perspective in motives and purposes: The case of adjunct clauses and theory of mind. Paper in symposium, complex language and theory of mind IASCL. July 18th, 2024.

Cheung, H. (2006). False belief and language comprehension in Cantonese-speaking children. *J. Exp. Child Psychol.* 95, 79–98. doi: 10.1016/j.jecp.2006.05.002

Cheung, H., Hsuan-Chih, C., Creed, N., Ng, L., Ping Wang, S. P., and Mo, L. (2004). Relative roles of general and complementation language in theory-of-mind development: evidence from Cantonese and English. *Child Dev.* 75, 1155–1170. doi: 10.1111/j.1467-8624.2004.00731.x

Cheung, H., Hsuan-Chih, C., and Yeung, W. (2009). Relations between mental verb and false-belief understanding in Cantonese-speaking children. *J. Exp. Child. Psych.* 104, 141–155. doi: 10.1016/j.jecp.2009.05.004

Cutting, A. L., and Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: individual differences and interrelations. *Child Dev.* 70, 853–865. doi: 10.1111/1467-8624.00061

Davidson, D. (1984). Thought and talk." inquiries into truth and interpretation. Oxford: Clarendon Press, 155–170.

De Mulder, H. N. M., Wijnen, F., and Coopmans, P. H. A. (2019). Interrelationships between theory of mind and language development: a longitudinal study of Dutch-speaking kindergartners. *Cogn. Dev.* 51, 67–82. doi: 10.1016/j.cogdev.2019.03.006

de Villiers, J. G. (1995). "Questioning minds and answering machines" in Proceedings of the Boston University conference on language development. eds. D. M. Laughlin and S. McEwen (Somerville, MA: Cascadilla Press).

de Villiers, J. G. (1999). "On acquiring the structural representations for false complements" in UMOP: New perspectives on language acquisition. ed. B. Hollebrandse (Amherst, MA: University of Massachusetts).

de Villiers, J. G. (2007). The interface of language and theory of mind. *Lingua* 117, 1858–1878. doi: 10.1016/j.lingua.2006.11.006

de Villiers, J. G. (2018). "Perspectives on truth: the case of language and false belief reasoning" in Semantics in language acquisition. eds. K. Syrett and S. Arunachalam (Amsterdam: John Benjamins), 222–245.

de Villiers, J. G. (2021). With language in mind. *Lang. Learn. Dev.* 17, 71–95. doi: 10.1080/15475441.2020.1820338

de Villiers, P. A., Burns, F., and Pearson, B. (2003). "The role of language in theory of mind development of language-impaired children: complementing theories" in Proceeding of the 27th annual Boston university conference on language development. eds. B. Beachley, A. Brown and F. Conlin, vol. *1* (Somerville, MA: Cascadilla Press).

de Villiers, J. G., and de Villiers, P. A. (2000). "Linguistic determinism and the understanding of false beliefs" in Children's reasoning and the mind. eds. P. Mitchell and K. Riggs (Hove: Psychology Press), 191–228.

de Villiers, J. G., and de Villiers, P. A. (2009). "Complements enable representation of the contents of false beliefs: the evolution of a theory of mind" in Language acquisition. ed. S. Foster (New York: Palgrave McMillian), 169–195.

de Villiers, P. A., and de Villiers, J. G. (2012). Deception dissociates from false belief reasoning in deaf children: implications for the implicit versus explicit theory of mind distinction. *Br. J. Dev. Psychol.* 30, 188–209. doi: 10.1111/j.2044-835X.2011.02072.x

de Villiers, J. G., and de Villiers, P. A. (2014). The role of language in theory of mind development. *Top. Lang. Disord.* 34, 313–328. doi: 10.1097/tld.0000000000000037

de Villiers, J. G., Harrington, E., Gadilauskas, E., and Roeper, T. (2012). "Tense and truth in children's question answering" in Proceedings of the Boston University conference on language development. eds. A. Biller, E. Chung and A. Kimball (Somerville, MA: Cascadilla Press).

de Villiers, J. G., and Pyers, J. E. (2002). Complement to cognition: a longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cogn. Dev.* 17, 1037–1060. doi: 10.1016/S0885-2014(02)00073-4

de Villiers, J. G., and Roeper, T. (2016). "The acquisition of complements" in Oxford handbook of developmental linguistics. eds. W. Snyder and J. Lidz (Oxford: Oxford University Press), 279–309.

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

Devine, R. T., and Hughes, C. (2012). Silent films and strange stories: theory of mind, gender, and social experiences in middle childhood. *Child Dev.* 84, 989–1003. doi: 10.1111/cdev.12017

Devine, R. T., and Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: a meta-analysis. *Child Dev.* 85, 1777–1794. doi: 10.1111/cdev.12237

Diessel, H., and Tomasello, M. (2001). The acquisition of finite complement clauses in English: a usage-based approach to the development of grammatical constructions. *Cogn. Linguist.* 12, 97–141. doi: 10.1515/cogl.12.2.97

Domitrovich, C. E., Cortes, R., and Greenberg, M. T. (2007). Improving young children's social and emotional competence: a randomized trial of the preschool PATHS curriculum. *J. Prim. Prev.* 28, 67–91. doi: 10.1007/s10935-007-0081-0

Durrleman, S., Burnel, M., de Villiers, J. G., Thommen, E., Yan, R., and Delage, H. (2019). The impact of grammar on mentalizing: a training study including children with autism spectrum disorder and developmental language disorder. *Front. Psychol.* 10, 1–15. doi: 10.3389/fpsyg.2019.02478

Durrleman, S., Burnel, M., Thommen, E., Foudon, N., Sonié, S., Reboul, A., et al. (2016). The language-cognition interface in ASD: complement sentences and false belief reasoning. *Res. Autism Spectrum Disord.* 21, 109–120. doi: 10.1016/j.rasd.2015.10.003

Durrleman, S., and de Villiers, J. G. (2024). Complements and false belief reasoning: An ecological training study in children with and without autism. Paper presented in symposium: "Complex language and theory of mind. Prague: International Association of Child Language.

Farrar, M. J., Benigno, J. P., Tompkins, V., and Gage, N. A. (2017). Are there different pathways to explicit false belief understanding? General language and complementation in typical and atypical children. *Cogn. Dev.* 43, 49–66. doi: 10.1016/j.cogdev.2017.02.005

Filippova, E., and Astington, J. W. (2010). Children's understanding of social-cognitive and social-communicative aspects of discourse irony. *Child Dev.* 81, 913–928. doi: 10.1111/j.1467-8624.2010.01442.x

Fizke, E., Butterfill, S., van de Loo, L., Reindl, E., and Rakoczy, H. (2017). Are there signature limits in early theory of mind? *J. Exp. Child Psychol.* 162, 209–224. doi: 10.1016/j.jecp.2017.05.005

Fodor, J. A. (1994). The human language series: Part 1: Discovering the human language. Director. G. Searchinger [video/DVD].

Farrant, B. M., Maybery, M. T., and Fletcher, J. (2012). Language, cognitive flexibility, and explicit false belief understanding: Longitudinal analysis in typical development and specific language impairment. *Child Dev*, 83, 223–235. doi: 10.1111/j.1467-8624.2011.01681.x

Grosse Wiesmann, C., Friederich, A. D., Singer, T., and Steinbeis, N. (2017). Implicit and explicit false belief development in preschool children. *Dev. Sci.* 20, 1–15. doi: 10.1111/desc.12445

Guo, Q., Pan, Q., Liu, Q., Wang, T., Cao, S., Lin, Y., et al. (2022). Relationship between different types of complement syntax and false belief in mandarin-speaking children with autism spectrum disorder and typically developing children. *Front. Psychol.* 13:1045227. doi: 10.3389/fpsyg.2022.1045227

Hale, C. M., and Tager-Flusberg, H. (2003). The influence of language on theory of mind: a training study. *Dev. Sci.* 6, 346–359. doi: 10.1111/1467-7687.00289

Harris, P. L., de Rosnay, M., and Pons, F. (2005). Language and children's understanding of mental states. *Curr. Directions Psychol. Sci.* 14, 69–73. doi: 10.1111/j.0963-7214.2005.00337.x

Heyes, C. (2014). False belief in infancy: a fresh look. *Develop. Sci.* 17, 647–659. doi: 10.1111/desc.12148

Hinzen, W. (2007). Mind design and minimal syntax. Oxford, UK: Oxford University Press.

Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Exp. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Jenkins, J. M., and Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Dev. Psychol.* 32, 70–78. doi: 10.1037/0012-1649.32.1.70

Klein, A., Starkey, P., and Ramirez, A. (2002). Pre-K mathematics curriculum. Glendale, IL: Scott Foresman.

Knuppel, A., Steensgaard, R., and Jensen de Lopez, K. (2007). Mental state talk by Danish preschool children. *Nordlyd* 34, 110–130. doi: 10.7557/12.125

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., and Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Dev.* 67, 490–507. doi: 10.2307/1131828

Korkman, M., Kirk, U., and Kemp, S. (1998). NESPY: A developmental neuropsychological assessment. San Antonio, TX: The Psychological Corporation.

Kuhnert, R. L., Begeer, S., Fink, E., and de Rosnay, M. (2017). Gender-differentiated effects of theory of mind, emotion understanding, and social preference on prosocial behavior development: a longitudinal study. *J. Exp. Child Psychol.* 154, 13–27. doi: 10.1016/j.jecp.2016.10.001

Kulke, L., and Rakoczy, H. (2018). Implicit theory of mind–an overview of current replications and non-replications. *Data Brief* 16, 101–104. doi: 10.1016/j.dib.2017.11.016

Kulke, L., von Duhn, B., Schneider, D., and Rakoczy, H. (2018). Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychol. Sci.* 29, 888–900. doi: 10.1177/0956797617747090

Leslie, A. M., German, T. P., and Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cogn. Psychol.* 50, 45–85. doi: 10.1016/j.cogpsych.2004.06.002

Li, H., and Leung, M. (2024). The roles of language and executive function in mandarin-speaking children's theory of mind development. *Front. Psychol.* 15:1354207. doi: 10.3389/fpsyg.2024.1354207

Lohmann, H., and Tomasello, M. (2003). The role of language in the development of false belief understanding: a training study. *Child Dev.* 74, 1130–1144. doi: 10.1111/1467-8624.00597

Lonigan, C. J., Clancy-Menchetti, J., Phillips, B. M., McDowell, K., and Farver, J. M. (2005). Literacy express: A preschool curriculum. Tallahassee, FL: Literacy Express.

Lonigan, C. J., Phillips, B. M., Clancy, J. L., Landry, S. H., Swank, P. R., Assel, M., et al. (2015). Impacts of a comprehensive school readiness curriculum for preschool children at risk for educational difficulties. *Child Dev.* 86, 1773–1793. doi: 10.1111/cdev.12460

Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: relations with complex syntactical mastery. *Child Dev.* 81, 597–615. doi: 10.1111/j.1467-8624.2009.01418.x

Low, J., and Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychol. Sci.* 24, 305–311. doi: 10.1177/0956797612451469

Mathieson, K., and Banerjee, R. (2011). Peer play, emotion understanding, and socio moral explanation: the role of gender. *Br. J. Dev. Psychol.* 29, 188–196. doi: 10.1111/j.2044-835X.2010.02020.x

Meins, E., Fernyhough, C., Wainwright, R., Das Gupta, M., Fradley, E., and Tuckey, M. (2002). Maternal mind-mindedness and attachment security as predictors of theory of mind understanding. *Child Dev.* 73, 1715–1726. doi: 10.1111/1467-8624.00501

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychol. Bull.* 135, 749–773. doi: 10.1037/a0016854

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

Mo, S., Su, Y., Sabbagh, M. A., and Xiu, J. (2014). Sentential complements and false belief understanding in Chinese mandarin-speaking preschoolers: a training study. *Cogn. Dev.* 29, 50–61. doi: 10.1016/j.cogdev.2013.11.001

Nelson, K. (2005). "Language pathways into the community of minds" in Why language matters for theory of mind. eds. J. W. Astington and J. A. Baird (Oxford: Oxford University Press), 26–49.

Nordmeyer, A. E., and de Villiers, J. G. (2019). To be or not to be: examining the role of language in a concept of negation. In Proceedings of the Annual Meeting of the Cognitive Science Society.

Onishi, K. H., and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* 308, 255–258. doi: 10.1126/science.1107621

Papafragou, A. (2001). "Linking early linguistic and conceptual capacities: the role of theory of mind" in Conceptual structure, discourse and language IV. eds. A. Cienki, B. Luka and M. Smith (Stanford, CA: CSLI Publications).

Perner, J., Leekam, S. R., and Wimmer, H. (1987). Three-year-olds' difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5, 125–137. doi: 10.1111/j.2044-835X.1987.tb01048.x

Perner, J., and Ruffman, T. (2005). Infants' insight into the mind: how deep? *Science* 308, 214–216. doi: 10.1126/science.1111656

Perner, J., Sprung, M., Zauner, P., and Haider, H. (2003). Want that is understood well before say that, think that, and false belief: a test of de Villiers' linguistic determinism on German-speaking children. *Child Dev.* 74, 179–188. doi: 10.1111/1467-8624.t01-1-00529

Poulin-Dubois, D., and Yott, J. (2018). Probing the depth of infants' theory of mind: disunity in performance across paradigms. *Dev. Sci.* 21:e12600. doi: 10.1111/desc.12600

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

Pyers, J. E. (2004). The relationship between language and false-belief understanding: Evidence from learners of an emerging sign language in Nicaragua. Unpublished doctoral dissertation. Berkeley: University of California.

Pyers, J. E., and Senghas, A. (2009). Language promotes false-belief understanding. *Psychol. Sci.* 20, 805–812. doi: 10.1111/j.1467-9280.2009.02377.x

Rakoczy, H. (2012). Do infants have a theory of mind? *Br. J. Dev. Psychol.* 30, 59–74. doi: 10.1111/j.2044-835X.2011.02061.x

Rakoczy, H. (2017). In defense of a developmental dogma: children acquire propositional attitude folk psychology around age 4. *Synthese* 194, 689–707. doi: 10.1007/s11229-015-0860-8

Reed, M. A., Pien, D. L., and Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Q.* 30, 131–147.

Repacholi, B. M., and Gopnik, A. (1997). Early reasoning about desires: evidence from 14-and 18-month-olds. *Dev. Psychol.* 33, 12–21. doi: 10.1037/0012-1649.33.1.12

Richard, M. (1990). Propositional attitudes: An essay on thoughts and how we ascribe them. Cambridge, U.K.: Cambridge University Press.

Ruffman, T., Slade, L., and Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Dev.* 73, 734–751. doi: 10.1111/1467-8624.00435

San Juan, V., and Astington, J. W. (2012). Bridging the gap between implicit and explicit understanding: how language development promotes the processing and representation of false belief. *Brit. J. Dev. Psychol.* 30, 105–122. doi: 10.1111/j.2044-835X.2011.02051.x

San Juan, V., and Astington, J. W. (2017). Does language matter for implicit theory of mind? The effects of epistemic verb training on implicit and explicit false-belief understanding. *Cogn. Dev.* 41, 19–32. doi: 10.1016/j.cogdev.2016.12.003

Scarborough, H. S. (1990). Index of productive syntax. *Appl. Psycholing* 11, 1–22. doi: 10.1017/S0142716400008262

Schick, B., de Villiers, P. A., de Villiers, J. G., and Hoffmeister, R. (2007). Language and theory of mind: a study of deaf children. *Child Dev.* 78, 376–396. doi: 10.1111/j.1467-8624.2007.01004.x

Schroeder, K., Durrleman, S., Çokal, D., Delgado, A. S., Marin, A. M., and Hinzen, W. (2021). Relations between intensionality, theory of mind and complex syntax in autism spectrum conditions and typical development. *Cogn. Dev.* 59:101071. doi: 10.1016/j.cogdev.2021.101071

Scott, R. M., and Baillargeon, R. (2017). Early false-belief understanding. *Trends Cogn. Sci.* 21, 237–249. doi: 10.1016/j.tics.2017.01.012

Segal, G. (1998). "Representing representations" in Language and thought. eds. P. Carruthers and J. Boucher (Cambridge: Cambridge University Press), 146–161.

Seymour, H., Roeper, T., and de Villiers, J. G. (2003). Diagnostic evaluation of language variation – screening test (DELV-ST). San Antonio, TX: Harcourt.

Shatz, M. (1994). "Theory of mind and the development of social-linguistic intelligence in early childhood" in Children's early understanding of mind: Origins and development. eds. C. Lewis and P. Mitchell (Hillsdale, NJ: Erlbaum).

Shatz, M., Wellman, H. M., and Silber, S. (1983). The acquisition of mental verbs: a systematic investigation of first references to mental state. *Cognition* 14, 301–321. doi: 10.1016/0010-0277(83)90008-2

Siu, C. T. S., and Cheung, H. (2022). A longitudinal reciprocal relation between theory of mind and language. *Cogn. Dev.* 62:101176. doi: 10.1016/j.cogdev.2022.101176

Southgate, V. (2013). "Early manifestations of mindreading" in Understanding other minds: Perspectives from developmental social neuroscience. eds. S. Baron-Cohen, M. Lombardo and H. Tager-Flusberg (Oxford: Oxford University Press), 3–18.

Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* 18, 587–592. doi: 10.1111/j.1467-9280.2007.01944.x

Southgate, V., and Vernetti, A. (2014). Belief-based action prediction in preverbal infants. *Cognition* 130, 1–10. doi: 10.1016/j.cognition.2013.08.008

Strachan, J. W., Albergo, D., Borghini, G., Pansardi, O., Scaliti, E., Gupta, S., et al. (2024). Testing theory of mind in large language models and humans. *Nat. Hum. Behav.* 8, 1285–1295. doi: 10.1038/s41562-024-01882-z

Tager-Flusberg, H., and Joseph, R. M. (2005). "How language facilitates the acquisition of false-belief understanding in children with autism" in Why language matters for theory of mind. eds. J. Astington and J. Baird (Oxford, U.K: Oxford University Press), 298–318.

Tardif, T., and Wellman, H. M. (2000). Acquisition of mental state language in mandarin-and Cantonese-speaking children. *Dev. Psychol.* 36, 25–43. doi: 10.1037/0012-1649.36.1.25

Tardif, T., Wing-Chee, S. C., and Kaciroti, N. (2007). Language and false belief: evidence for general, not specific, effects in Cantonese-speaking preschoolers. *Dev. Psychol.* 43, 318–340. doi: 10.1037/0012-1649.43.2.318

Tomasello, M. (1999). The cultural origins of human cognition. Cambridge, MA: Harvard University Press.

Tompkins, V., Benigno, J., Kiger Lee, B., and Wright, B. (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

Wagner, R. K., Torgesen, J. K., and Rashotte, C. A. (1999). Comprehensive test of phonological processing (CTOPP). Austin, TX: Pro-Ed.

Walker, S. (2005). Gender differences in the relationship between young children's peer-related social competence and individual differences in theory of mind. *J. Genet. Psychol.* 166, 297–312. doi: 10.3200/GNTP.166.3.297-312

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 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. (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

Westra, E., and Nagel, J. (2021). Mindreading in conversation. *Cognition* 210:104618. doi: 10.1016/j.cognition.2021.104618

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

Winner, E. (1997). The point of words: Children's understanding of metaphor and irony. Cambridge, MA: Harvard University Press.

Woodward, A. L. (2009). Infants' grasp of others' intentions. *Curr. Dir. Psychol. Sci.* 18, 53–57. doi: 10.1111/j.1467-8721.2009.01605.x

Woolfe, T., Want, S. C., and Siegal, M. (2002). Signposts to development: theory of mind in deaf children. *Child Dev.* 73, 768–778. doi: 10.1111/1467-8624.00437

Yu, C., and Wellman, H. M. (2024). A meta-analysis of sequences in theory-of-mind understandings: theory of mind scale findings across different cultural contexts. *Dev. Rev.* 74:101162. doi: 10.1016/j.dr.2024.101162

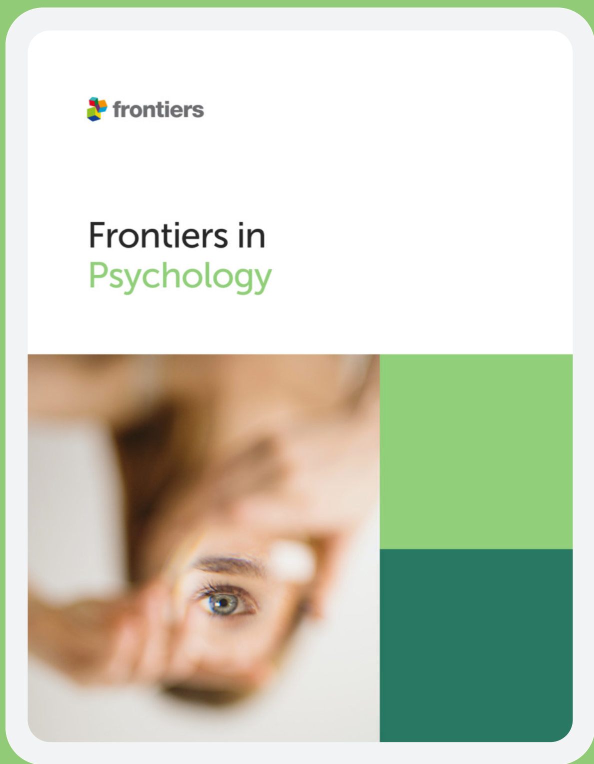
Frontiers in Psychology

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