

Theory of Mind in Preadolescents: Assessing the Effects of a Conversational Training

Virgina Rinaldi

A.B.C. Associazione per i Bambini Chirurgici del Burlo OdV ETS

Claudia Leone

Department of Life Sciences, University of Trieste

Angela Camelio

A.B.C. Associazione per i Bambini Chirurgici del Burlo OdV ETS

Marcella Caputi and Alessandro Cuder

Department of Life Sciences, University of Trieste

Giuseppina Battain

A.B.C. Associazione per i Bambini Chirurgici del Burlo OdV ETS

Sandra Pellizzoni

Department of Life Sciences, University of Trieste

Studies on theory of mind (ToM) have traditionally focused on preschool age, neglecting preadolescence and adolescence. However, ToM skills are equally fundamental in these age groups given the continued development of mental abilities and their link to social success. The aim of this study is to measure any changes in ToM skills in a group of middle school students who participated in three intervention sessions that followed a conversational approach focused on perspective-taking and the inclusion of hospitalized children. This investigation is especially relevant after COVID-19 and its consequences among adolescents and preadolescents. The intervention had positive effects on students' ToM levels, improving their ability to understand intentions and attribute mental states to other people. Future studies could provide a more comprehensive picture of the results by examining other aspects of adolescent socio-cognitive functioning. Data are discussed in terms of promoting well-being during this developmental period.

Keywords: theory of mind, preadolescence, conversational training

INTRODUCTION

Theory of mind (ToM) plays a central role in research on how individuals attribute mental states to themselves and others, shaping social, cognitive, and emotional development across the lifespan (Carlson et al., 2013). As noted in a recent commentary by Quesque and colleagues (2024), different definitions of ToM coexist in the literature, leading to the imprecise grouping of different aspects of mental state attribution under a single label (Warnell & Redcay, 2019; Yeung et al., 2024). According to these authors, ToM is typically defined as the ability to infer others' mental states, an ability they suggest is more accurately described by the more general label "mentalizing" (Quesque et al., 2024). More specifically, they propose that ToM should be understood as "the use of folk psychological knowledge and heuristics to think about one's own and other people's mental states," a definition that more closely reflects the traditional conception of ToM and points to a specific way of mentalizing (Quesque et al., 2024, p. 2). ToM, which in turn, appears to comprise distinct components: socio-perceptual components, which involve decoding observable cues (e.g., facial expressions, gestures), and socio-cognitive components, which involve inferring what others think, believe, or intend in the absence of observable information (Meinhardt-Injac et al., 2020; Tager-Flusberg & Sullivan, 2000).

Studies have shown that ToM is essential for various communicative and cognitive functions, including conversation (e.g., Grice, 1989; Sperber & Wilson, 1995), understanding criticism (Cutting & Dunn, 2002), deception (Sodian, 1991), humor and lying (Hughes & Leekam, 2004; Leekam & Prior, 1994), and irony (Happé, 1994). Moreover, ToM has been linked to pragmatic language competence (Eisenmajer & Prior, 1991) and problem solving (Greenberg et al., 2013). ToM also plays an important role in peer interactions: its preschool levels lead to peer acceptance/rejection in the early primary school years via prosocial behavior (Caputi et al., 2012). Moreover, mature ToM skills can contribute to reduce levels of social anxiety via social acceptance during middle school years (Ronchi et al., 2020).

Many studies investigated the development of ToM in early childhood and showed that early ToM abilities are predicted by language development, cognitive functions, socio-demographic factors, and neural correlates (Diaz & Farrar, 2018; Navarro & Conway, 2021; Navarro et al., 2021; Rakoczy, 2022; Rubio-Fernández & Glucksberg, 2012). However, the processes underlying ToM development remain unclear, partly due to the variability in how researchers operationalize the construct and the wide range of tasks used to assess it (Navarro & Conway, 2021). Schaafsma and colleagues (2015) suggested that one solution to this issue is to avoid treating ToM as a "monolithic" ability. Instead, researchers should consider the flexible nature of the construct. The theoretical mechanisms involved in the development of ToM are the subject of numerous recent studies that examine the findings from social cognition, neuroscience, and child development fields. The main theoretical mechanisms focus on the interaction between cognitive and emotional skills, the integration of mental representations, and the maturation of brain systems (Apperly, 2012; Beauchamp & Anderson, 2010; Mar, 2011). Other

research investigated possible processes relevant to the development of ToM in healthy populations. These processes interact with various cognitive, emotional, and environmental aspects, promoting a gradual understanding of others' mental states (for a review see Rakoczy, 2022). For instance, through symbolic play, children may learn to understand that other people can have emotions, desires, and beliefs different from their own (Harris, 1992). Furthermore, children who grow up in a loving environment, where they can freely explore their own emotions, develop a better ability to understand others (Denham, 2006). Similarly, interactions with adults and peers help develop an understanding of the distinction between one's own point of view and that of others (Wood & Middleton, 1975), as well as empathy and cooperation (Greenfield, 2000). Different studies suggest that ToM skills continue to develop during middle childhood, adolescence, and adulthood (Dumontheil et al., 2010; Weimer et al., 2017). In preadolescence, specifically, the understanding of others' mental states (Devine & Hughes, 2013) and the ability to reason about ambiguous social interactions (Bosacki, 2000) both continue to improve. In this context, Advanced Theory of Mind or aToM (Białycka-Pikul et al., 2017) refers to more sophisticated and complex ToM abilities (evaluated in subjects older than preschoolers) such as, for example, the ability to recognize when people are pretending. Adolescence involves the need to deduce complex social emotions based on beliefs, which are deeply rooted in social contexts and the life stories of others (Hoffman, 2001). Given the distinctive nature of adolescents' social relationships and their need to navigate increasingly complex interpersonal situations (Blakemore, 2008; Crone & Dahl, 2012), this developmental stage may play a crucial role in shaping young people's aToM abilities (Hughes, 2011), underscoring the importance of further investigation into this period of development.

ToM abilities are fundamental for understanding diversity and fostering inclusiveness. When people are perceived as "other," the spontaneous tendency to mentalize decreases; conversely, perspective-taking reduces stereotypes and bias (Fiske, 2009; Galinsky & Moskowitz, 2000; Hoyo et al., 2019). Attributing mental states to and considering the perspectives of individuals from different groups (e.g., peers with disabilities, those from ethnic-minority backgrounds, or children who are hospitalized) promotes the social inclusion of these groups (Vezzali et al., 2012) while simultaneously strengthening ToM competencies of those who engage in these activities (Bianco et al., 2016; Ornaghi et al., 2014). Therefore, it is important to take immediate action by fostering skills such as valuing diversity and, consequently, ToM skills, such as recognizing others' emotions and considering different perspectives, while avoiding stereotypes and prejudices (Nota et al., 2015). However, only few studies so far investigated the effectiveness of school-based trainings in promoting these aspects in middle schools. Hence, the present study aims to evaluate the effectiveness of a conversational intervention, focused on perspective-taking and the inclusion of hospitalized children, in promoting socio-cognitive and socio-perceptual ToM abilities in a sample of middle-school students.

TOM TRAININGS IN EXISTING LITERATURE

In order to place our study within the existing literature, it is worth reviewing studies which propose ToM trainings. Most training programs to promote ToM skills targeted preschoolers (e.g., Honoré et al., 2020; Lecce, Caputi, & Pagnin, 2014) or clinical groups, like children with autism or hearing impairment (Berger et al., 2017; Wellman & Peterson, 2013). Interventions targeting preadolescents are scarce, and trainings that foster ToM in typically developing adolescents and adults are even rarer. Bianco and Castelli (2023) note that key facets of mature ToM remain largely unexplored in adolescents and adults' populations. Yet the few existing studies show that strengthening these skills enhances mentalizing abilities (Lombardi et al., 2022), social functioning (Filippello et al., 2022), emotion regulation, self-control, and resilience (Stein, 2006; Lombardi et al., 2022). Expanding research on ToM in adolescence and adulthood is therefore a priority.

Considering the existing literature, meta-analytic evidence (Hofmann et al., 2016) based on participants aged 2 to 16 indicates that interventions are effective in enhancing ToM skills. The authors also highlight that methods involving active engagement and encouraging discussion about one's own and others' mental states such as imaginative play, modeling, role-playing, and corrective feedback are particularly effective. Similarly, some researchers showed that conversations about others' mental states are effective in enhancing ToM skills (Bianco et al., 2016; Ornaghi et al., 2014). According to the conversational approach, understanding mental states develops through dialogue in children: talking about the mind fosters coordination of different views on the same event, stimulates comparison of one's own mental states with those of others, and promotes reflection on social experiences (Carpendale & Lewis, 2004; Nelson, 2005; Symons, 2004; Turnbull & Carpendale, 1999). Moreover, the effectiveness of the conversational approach is amplified when it is accompanied by feedback and explanatory scaffolding (Clements et al., 2000; Melot & Angeard, 2003). However, intervention studies targeting ToM skills in samples aged 11 to 13 are still scarce, particularly those employing a conversational approach, whereas research involving younger children (under 11 years old) is more common. For example, Lecce, Bianco, and colleagues (2014), in a sample of 9-year-olds, found that a ToM intervention based on a conversational approach was effective in enhancing ToM skills, with effects maintained at a two-month follow-up. Similarly, Bianco and colleagues (2016) reported positive outcomes for ToM through conversations about mental states in peer interactions. In line with these findings, more recent studies (Caputi et al., 2021; Gao et al., 2020; Valle et al., 2016) have also shown the effectiveness of conversational approaches in promoting ToM development. In addition, different strategies have been proposed to foster ToM skills during conversational interventions, especially considering older children. One is the think-aloud method, in which students express their thoughts while reading stories or observing others, thereby fostering the comprehension of others' mental states (Beck & McKeown, 2006). Another strategy to foster ToM skills is role-playing (Jha & Sharma, 2020): by acting out real or fictional

scenarios, adolescents may practice recognizing emotions, non-verbal cues, and the perspectives of all characters involved, skills that become increasingly important as their social worlds grow more complex.

THE PRESENT STUDY

This study, conducted in collaboration with the University of Trieste, Department of Life Sciences, is part of a broader project named “ABC Model” promoted by ABC (Associazione per i Bambini Chirurgici del Burlo), the association connected to the Children’s Hospital Burlo Garofolo, Surgery Department, in Trieste. The ultimate goal of the dedicated project “A scuola con ABC: educazione e formazione alla cultura del volontariato e dell’inclusione” (At School With ABC: Education and Training in the Culture of Volunteering and Inclusion) was the promotion of school inclusion and diversity, in order to contribute to the development of awareness in young citizens. Indeed, as outlined in the Introduction, research indicates that conversational activities centered on themes of diversity, otherness, and inclusion not only help reduce exclusionary dynamics toward classmates perceived as “different” (e.g., due to ethnic background, chronic illness, or hospitalization), but also foster the development of students’ ToM skills. Promoting interventions that frame diversity as a resource and actively engage students in socio-cognitive processes is therefore essential within the school context. To pursue the goal of this project, we conducted a study that considered the effects of a conversational approach intervention on ToM development. This focus is relevant during pre-adolescents and adolescents (Blakemore, 2008; Crone & Dahl, 2012; Hughes, 2011), given the limited number of studies that investigated the effectiveness of school-based training programs aimed at fostering ToM, especially in middle school.

The aim of the present study was to assess the effectiveness of an intervention in promoting ToM in middle school students. In doing so, the study examined two groups of students attending middle school: an intervention group and a passive control group. Three training sessions led by a researcher from the association were conducted with the intervention group. These sessions included the presentation and reading of stories, videos, discussions, and comparisons regarding various issues. The topics were selected with consideration for the fact that the target cohort had experienced limited social interactions due to COVID-19 restrictions over the previous two years (Deolmi & Pisani, 2020; Sabato et al., 2021). The intervention consisted of conversational activities focused on understanding others’ mental states and individual differences, including discussions around the experiences of hospitalized children, reflections on videos and excerpts, role-playing exercises exploring the themes of otherness and inclusion, and activities aimed at fostering altruism and promoting the value of volunteering. Evaluation of the effectiveness of the intervention was conducted using three different types of ToM tests before the three intervention training sessions (pre-test) and afterwards (post-test). A verbal reasoning test was used as a control variable to compare the groups at pre-test, as a good knowledge of the Italian language was essential for understanding the interventions delivered in the classroom by the researcher. Overall,

numerous studies suggest that ToM abilities continue to develop through middle childhood, adolescence, and into adulthood (Dumontheil et al., 2010; Weimer et al., 2017). From a developmental perspective, adolescence is viewed as a critical period during which ToM interventions are both needed and potentially highly effective. Theoretically, ToM can be seen as a developing capacity with important milestones emerging during pre-adolescence and adolescence. This stage is marked by rapid cognitive growth (e.g., executive functions), emotional development, and increasing social complexity. Unlike earlier childhood, however, this developmental surge coincides with heightened sensation seeking and still-maturing self-regulation, a combination that can increase risk-taking behaviors, especially in peer contexts (Chein et al., 2011; Romeo, 2013). Moreover, adolescents are particularly sensitive to peer judgment, often experiencing challenges in interpersonal relationships and a heightened sense of exclusion. In light of these developmental features, we designed activities that focused on perspective-taking, dialogic interaction, and small-group reasoning, with the goal of supporting students' cognitive, emotional, and social skills. Previous literature testifies the effectiveness of such approaches in promoting ToM in this age group (Bianco et al., 2016; Ornaghi et al., 2014). The hypothesis of the present study was that greater gains in ToM skills would be observed in the intervention group than in the control group at the post-test. Specifically, we hypothesized that the students would demonstrate improvement from pre-test to post-test in the levels of ToM assessed by the False Belief and Faux Pas Tests, that is, in the aspects that are more related to socio-cognitive components (Meinhardt-Injac et al., 2018). We did not expect any improvement in the Eyes Test in light of the fact that we did not propose activities related to the socio-perceptual aspects of ToM (Meinhardt-Injac et al., 2020).

MATERIALS AND METHODS

PARTICIPANTS

Before conducting the study, a power analysis using G*Power (Faul et al., 2007) was performed to determine the minimum number of participants required. Assuming the use of an independent sample *t*-test, hypothesizing a medium difference between the groups, with a critical alpha of .05, power of 0.80, and a group allocation ratio of 0.25, the power analysis recommended recruiting 200 participants. We then used a regional database to randomly contact school principals and teachers to present the project and propose the planned activities. Upon receiving approval from both the principals and the teachers, and in accordance with the procedures outlined in the Declaration of Helsinki and the guidelines of the Italian Association of Psychology, written informed consent was obtained from participants' parents or legal guardians. Participation was voluntary, with school involvement based on the decision of the principal and teachers, and student participation contingent on parental consent. In addition, verbal assent was sought from the students before their participation in the activities. A total of 233 participants were contacted. Nonetheless, 11 were excluded for the

following reasons: three were absent during the pre- and post-test evaluation, six were absent on two or more meetings, and two had a neurodevelopmental disorder that specifically affects their theory of mind. Therefore, 222 participants ($M_{\text{age}} = 12.28$, $SD_{\text{age}} = 0.73$, females = 115, males = 107) attending middle schools in Trieste took part in the study. Participants' school classes were randomized in two groups: an intervention group (classes = 12; $n = 170$, $M_{\text{age}} = 12.38$ years, $SD_{\text{age}} = 0.80$; females = 83, males = 87) and a passive control group (classes = 3; $n = 52$, $M_{\text{age}} = 11.98$ years, $SD_{\text{age}} = 0.24$, females = 32, males = 20). The students attended public schools located in the same city in northeastern Italy. Each class was supervised by a different teacher, who had been instructed prior to the start of the project to continue with their usual teaching activities. In the Italian school system, the middle school curriculum is standardized nationwide and defined by the Ministry of Education.

PROCEDURES

The study was structured into three phases: a pre-test phase, an intervention phase for the intervention group, and a post-test phase (Figure 1). All assessments and intervention sessions were conducted by the same researcher, who is also a licensed professional psychologist. Both the assessments and the intervention took place during school hours in the spring term, at times scheduled in coordination with the classroom teacher. All activities were carried out in the students' regular classroom, where they typically attend their lessons.

The pre-test evaluation lasted approximately 45 minutes and was conducted collectively in every classroom. The tests were presented in an abbreviated version due to time constraints. The researcher read the instructions of each test aloud and invited students to ask any questions before beginning. To minimize peer influence and conformity effects, students were instructed to complete the tests in silence, and seating was arranged to discourage interaction or copying. All participants completed paper-and-pencil assessments measuring verbal reasoning and ToM abilities (i.e., the Faux Pas Test, the False Belief Test, and the Eyes Test; see the Appendix for item examples). One week later, the intervention group began the program, which consisted of three one-hour sessions held once a week (a total of three hours of intervention). The control group did not take part in the intervention and continued with regular school schedule, participating only in the two evaluation phases (i.e., pre-test and post-test assessments). Two weeks after the third (final) intervention session, the post-test evaluation was administered collectively and lasted approximately 35 minutes. It included the same ToM measures used in the pre-test (the verbal reasoning test was not readministered).

INTERVENTION

The intervention consisted of three training sessions led by a researcher, featuring engaging activities centered around three main themes: (1) diversity and hospitalization, (2) otherness and inclusion, and (3) altruism and volunteering. In the

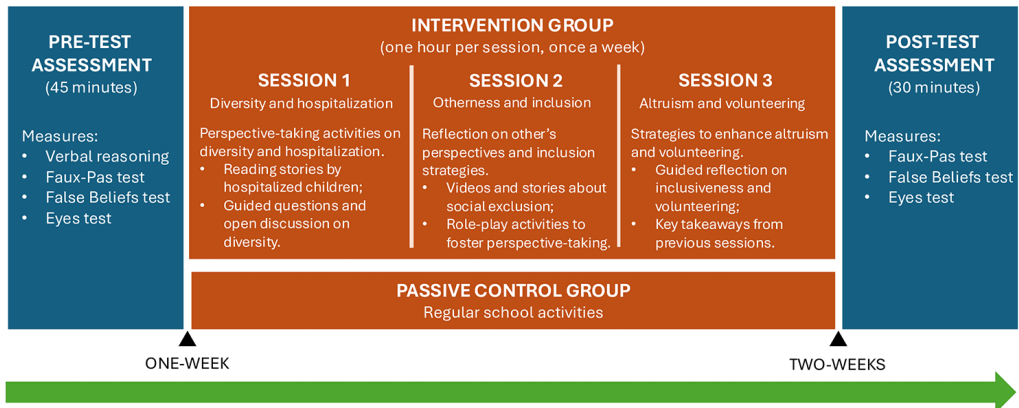


FIGURE 1. Overview of the study design, including the assessment phases and the structure of the three-session intervention program.

first session, following a brief self-introduction, the researcher presented the work of the ABC Association, an organization that supports children undergoing long-term hospital care and their parents. The session continued with the reading of a series of first-person stories written by hospitalized children. After each story, the class paused to reflect on the protagonists' thoughts, feelings, and everyday challenges (e.g., keeping up with schoolwork, coping with physical barriers). Guided questions encouraged students to compare their own daily experiences with those of the children in the stories and to infer the mental states behind their behaviors. The session concluded with an open discussion, during which students shared their reflections, and the researcher helped them make explicit connections to perspective-taking in everyday life.

In the second session, the activities focused on recognizing otherness and transforming understanding into inclusive actions. Students were shown short videos and read excerpts that illustrated social exclusion related to disability or medical conditions. They then participated in two role-play scenarios: one set in a pediatric ward and the other in a classroom (for an example see the Appendix). Each student was assigned a scripted role (e.g., patient, classmate, nurse, teacher) and engaged in interactive activities under the researcher's supervision. After the role-play, the researcher facilitated a debriefing session, emphasizing how recognizing another person's perspective helps to understand not only what someone does, but why they do it, by reflecting on their mental states. As a group, students then brainstormed practical strategies to make school, sports, and recreational settings more inclusive.

In the final session, the researcher introduced practical ways to turn perspective-taking into inclusive behaviors, with a particular emphasis on altruism and volunteering. Students began by reviewing key insights from the previous sessions and were then invited to consider how acts of altruism could help address the

everyday difficulties faced by hospitalized children. This reflection took place individually, supported by guiding questions from the researcher, and was followed by a class discussion in which students shared their ideas for supporting peers in similar situations. To underscore the conceptual link between altruism and the act of giving, each participant was provided with an object (e.g., a canvas tote bag) to personalize as part of the final activity. This creative task was designed not only to symbolize altruistic behavior, but also to offer a visual and tangible representation of the value of their own diversity and individuality, thereby deepening the students' engagement with the themes of uniqueness and inclusion.

To close the intervention, the researcher provided a summary of the three sessions, highlighting the key points and reinforcing the message that understanding others' mental states can lead to meaningful and inclusive behaviors.

MEASURES

Verbal Reasoning Test. A reduced version of the Verbal Reasoning Test from the battery of Primary Mental Abilities (PMA; see Thurstone, 1962), was used to test students' lexical skills and vocabulary. The test comprised 25-word items, each asking the student to identify the respective synonym among five proposed alternatives. The score was represented by the total number of correct answers (1 point for each correct answer and 0 points for each incorrect answer). The minimum possible score was 0 points and the maximum score obtainable is 25 points. The Verbal Reasoning Test was employed to evaluate potential differences between the intervention and passive control groups at pre-test. The reliability for the task was good in the present study (McDonald's omega, $\omega = 0.81$).

Faux Pas Test. The Faux Pas Test (Stone et al., 1998) was used in its Italian version (adapted by Liverta-Sempio et al., 2005), with 3 out of 10 faux pas presented in the version for adults. Each item consists of a series of social gaffes that occur when someone says something they should not have said (Gregory et al., 2002). Understanding a faux pas requires at least two distinct mental state representations, making it an advanced ToM task (Li et al., 2013; Shamay-Tsoory et al., 2010): (1) a cognitive ToM representation of the speaker's belief (i.e., the speaker's unawareness that the remark was inappropriate); and (2) an affective ToM representation of the listener's emotional state, as the listener feels insulted or hurt (Stone et al., 2003). The test consists of three stories. The questions which follow these stories are designed to determine whether the pupils correctly identified the presence of social error. For each story, participants first answered a screening question asking whether a faux pas had occurred. If they failed to detect it, the story received a score of 0. If they detected it, they proceeded to five follow-up questions probing their understanding of why the remark was inappropriate, the speaker's beliefs, and the listener's feelings. Five of these six questions were scored (1 = correct, 0 = incorrect), yielding a maximum of 5 points per story and a total score ranging from 0 to 15 across the three stories. The reliability observed in the present study for the task was good (McDonald's omega, $\omega = 0.94$).

False Belief Test. The False Belief Test (Confalonieri, 2009) was used to investigate the ability to ascribe beliefs, a fundamental aspect of ToM. Stories are read aloud one at a time, and then time is allowed for students to individually answer questions that investigate second- and third-order beliefs. Each question requiring a right or wrong answer is always followed by a justification question which asks for an explanation of the given answer to make the steps in mental reasoning explicit. We selected two stories (one for second order and one for the third) out of six proposed in the original task. One point was awarded for each correct answer. Scores ranged from 0 to 4 points for the first story and from 0 to 5 points for the second. The scores from both stories were then summed to obtain a total score ranging from 0 to 9 points. The reliability observed in the present study for the task was good (McDonald's omega, $\omega = 0.89$).

The Eyes Test. The Eyes Test was the final test to be administered (Baron-Cohen et al., 2001; Bolte, 2005; Serafin & Surian, 2004). This test, whose complete name is Reading the Mind in the Eyes Test, assesses the ability to attribute complex mental states (such as an emotion) in adults and adolescents (Baron-Cohen et al., 2001). Students are asked to carefully observe 15 photographs of gazes on a screen for three seconds, and to indicate the term that best describes the gaze, choosing among four possible alternatives. They are awarded 1 point for each correct answer and 0 points for each incorrect answer. The minimum score is 0 points, and the maximum obtainable score is 15 points. The reliability for the task in the present study was scarce (McDonald's omega, $\omega = 0.35$).

RESULTS

Data analysis was conducted using R statistical software. The data were analyzed for 222 participants (i.e., 170 participants in the intervention group and 52 in the passive control group). However, prior to the analysis, one student was excluded due to outlier performance on the PMA test and the corresponding data point was removed using listwise deletion. Multiple imputation was used to handle missing data, employing the predictive mean matching method to replace missing values with plausible estimates. Five imputed datasets were generated and used to obtain pooled regression estimates (see Van Buuren & Groothuis-Oudshoorn, 2011). In order to assess the improvements in pupils' ToM skills, comparisons were made between the intervention and control groups by calculating pre-post differences for each variable score. Table 1 shows the descriptive statistics for each group and each instrument used in the two phases of assessment (pre-test and post-test assessments).

PRE-TEST ANALYSIS

To assess the potential non-independence of observations due to classroom-level clustering, we calculated intraclass correlation coefficients (ICCs) for all dependent variables. ICCs ranged from 0.061 to 0.086, indicating a modest clustering

TABLE 1. Descriptive Statistics for the Two Groups, Including Raw Scores and Gain Scores for the Variables Considered in the Study

Variables	Intervention group							Passive control group						
	<i>n</i>	<i>M</i>	<i>SD</i>	Min-Max	Sk.	Ku.	<i>SE</i>	<i>n</i>	<i>M</i>	<i>SD</i>	Min-Max	Sk.	Ku.	<i>SE</i>
Raw scores														
PMA	169	17.04	3.08	7–23	–0.58	0.25	0.24	52	16.52	3.32	9–23	–0.31	–0.65	0.46
Faux Pas Test (T1)	170	7.09	3.62	0–14	–0.21	–0.84	0.28	52	8.04	3.42	0–14	–0.66	0.01	0.47
False Beliefs Test (T1)	170	7.24	1.61	2–9	–0.84	0.28	0.12	52	7.35	1.47	2–9	–1.00	1.43	0.20
Eyes Test (T1)	170	7.45	2.22	3–13	0.20	–0.40	0.17	52	6.90	2.05	4–11	0.23	–1.01	0.28
Faux Pas Test (T2)	170	7.29	4.11	0–15	–0.16	–1.08	0.32	52	8.06	3.37	0–14	–0.67	0.08	0.47
False Beliefs Test (T2)	170	7.59	1.68	0–9	–1.55	2.82	0.13	52	7.33	1.50	2–9	–0.89	1.01	0.21
Eyes Test (T2)	170	7.18	2.23	2–13	–0.18	–0.55	0.17	52	6.87	1.99	4–11	0.28	–0.86	0.28
Gain scores														
Faux Pas Test	170	0.19	3.19	–9–8	–0.34	0.59	0.24	52	0.02	0.31	–1–1	0.45	6.95	0.04
False Beliefs Test	170	0.35	1.60	–4–5	0.29	0.64	0.12	52	–0.02	0.37	–1–1	–0.23	4.12	0.05
Eyes Test	170	–0.27	2.42	–6–7	0.06	0.42	0.19	52	–0.04	0.34	–1–1	–0.64	5.19	0.05

Note. *M* = mean, *SD* = standard deviation, Min-Max = minimum and maximum score; Sk. = skewness, Ku. = kurtosis, *SE* = standard error.

effect. For this reason, we proceeded with the analyses without employing multilevel modeling to account for the nested structure of the data. In order to assess whether significant differences existed between the two groups at pre-test, Welch’s test was used. With unequal sample sizes, Student’s t-test can inflate Type I error in the presence of heteroscedasticity. Welch’s test controls this risk and yields comparable power when variances are equal (Delacre et al., 2017; Ruxton, 2006). Specifically, each measure (i.e., PMA, Faux Pas, False Beliefs, Eyes Test) was set as the dependent variable and the group factor (i.e., intervention and control group) was set as the independent variable. Hedges’s g^* was used to report effect sizes because it appears to be a better estimator in the case of violation of the assumption of homoscedasticity of variances (Delacre et al., 2021). In the following analyses, we also report the differences in group means, calculated by subtracting the control group mean from the intervention group mean (i.e., referenced as $M_{\text{intervention} - \text{control}}$).

The results showed that the two groups at pre-test did not differ from each other in the PMA, $F(1, 79.93) = 1.019, p = .316$, Hedges’s $g^* = 0.17$, $M_{\text{intervention} - \text{control}} = 0.52$; in the Faux Pas Test, $F(1, 88.80) = 2.956, p = .089$, Hedges’s $g^* = 0.26$, $M_{\text{intervention} - \text{control}} = -0.95$; in the False Beliefs Test, $F(1, 91.98) = 0.217, p = .642$, Hedges’s $g^* = 0.07$,

$M_{\text{intervention} - \text{control}} = -0.11$; and in the Eyes Test, $F(1, 90.75) = 2.683$, $p = .105$, Hedges's $g^* = 0.25$, $M_{\text{intervention} - \text{control}} = 0.55$. There were no statistically significant differences in gender distribution between the groups, $\chi^2(1) = 2.094$, $p = .148$.

POST-TEST ANALYSIS

In order to assess the gain at post-test compared with pre-test scores, we calculated the gains for each variable (GAIN = post-test – pre-test). This analytical strategy is commonly used in developmental age training studies (Alloway et al., 2013; Brehmer et al., 2012; Cuder et al., 2022). We then conducted Welch's tests by placing relative gains on each measure as the dependent variable and on the group factor as the independent variable.

For the Faux Pas Test, the gain analysis of the two groups conducted by Welch's test showed that the intervention group did not differ statistically from the control group, $F(1, 179.2) = 0.496$, $p = .482$, Hedges's $g^* = 0.06$, $M_{\text{intervention} - \text{control}} = 0.17$. This result indicates that the children, following the intervention, did not improve their understanding of mental states assessed with this instrument.

For the False Belief Test, Welch's test showed that the intervention group showed higher gains than the control group $F(1, 211.79) = 7.827$, $p = .005$, Hedges's $g^* = 0.26$, $M_{\text{intervention} - \text{control}} = 0.37$. This result indicates an improvement in the ability to attribute and understand another person's beliefs and intentions following the intervention.

For the Eyes Test, the gain analysis of the two groups conducted by Welch's test showed that the intervention group did not statistically differ from the control group, $F(1, 188.99) = 1.469$, $p = .227$, Hedges's $g^* = 0.11$, $M_{\text{intervention} - \text{control}} = 0.23$. This indicates that subjects in the intervention group did not improve their ability to attribute emotions to others by gaze expression.

DISCUSSION

The present study is part of a project promoted by ABC Association which aims to raise awareness of diversity and school inclusion. As mentioned in previous sections, extant literature shows that ToM plays a significant role in individual development, particularly in the context of social interactions among students (Caputi et al., 2012; Ronchi et al., 2020). This study, conducted in collaboration with the University of Trieste, Department of Life Sciences, evaluated the effects of an intervention program conducted with middle school students through a conversational approach concerning inclusion of hospitalized children.

Specifically, we expected that the intervention would improve the performance in two different socio-cognitive measures of ToM, i.e., the Faux Pas Test and False Belief Test. The results showed that, in the False Belief Test, the intervention group obtained statistically significant higher gains than the control group. Thus, students who took part in the intervention showed an improvement in their ability to attribute and understand another person's beliefs and intentions, compared with students belonging to the control group, who did not take part in the training

sessions. This result aligns with previous evidence showing the effectiveness of conversational interventions in promoting socio-cognitive components of ToM (Bianco et al., 2015; Caputi et al., 2021; Hofmann et al., 2016). However, we did not find statistically significant differences in the gains of the two groups in the Faux Pas Test. One likely reason is the reduced sensitivity of the abbreviated, three-item version of the Faux Pas Test that we employed, which may have masked more subtle post-intervention gains. This finding must also be interpreted in light of the well-known difficulty of measuring ToM with different instruments. Numerous studies have shown that commonly used ToM tasks are often only weakly correlated with one another (e.g., Warnell & Redcay, 2019), suggesting that they capture distinct facets of an inherently complex construct. Consistent with this view, we observed a modest pre-test correlation between the False Belief and the Faux Pas Test in our study ($r = .24, p < .01$). This supports the idea that improvement in one facet of ToM does not necessarily entail improvement in another. Future research should replicate our findings and examine more closely which specific aspects of ToM are most likely to improve through conversational interventions.

No statistically significant group differences were found on the Eyes Test. This may be due to the fact that the intervention did not explicitly target the decoding of mental states from non-verbal cues; such skill is also shaped by factors like personal experience, empathy, and sensitivity to social dynamics (Serafin & Surian, 2004). In addition, the Eyes Test itself shows relatively low internal consistency (Baron-Cohen et al., 2001; Dehning et al., 2012), which may have reduced its ability to detect subtle gains. Indeed, recent studies show psychometric limitations of this test. In a meta-analysis, Kittel and colleagues (2021) reported modest internal consistency and a weak factorial structure, suggesting that the Eyes Test operates more as a measure of emotion perception rather than ToM. Similarly, Higgins and colleagues (2022) recommend exercising caution when using the Eyes Test as a measure of ToM, citing its poor psychometric properties and the low inter-item correlations observed in their study. More recently, Higgins and colleagues (2024) found that the majority of studies employing the Eyes Test did not report evidence of construct validity, and when such evidence was provided, it often failed to meet accepted psychometric standards, raising concerns about the interpretability and validity of the resulting scores.

In conclusion, the results of the present study suggest that an intervention stimulating reflection on the inclusion of hospitalized children, including perspective-taking activities, teaching strategies to foster altruism, may enhance certain (socio-cognitive) components of ToM in middle school students, particularly their ability to understand intentions and attribute mental states to others. However, the findings were less conclusive for the Faux Pas Test and the Eyes Test, suggesting that future studies should consider carrying out a higher number of intervention sessions and employing more reliable assessment tools. These results partially align with previous training studies conducted with typically developing school-aged populations and provide preliminary evidence of effectiveness in an age group that remains relatively underrepresented in the literature (Lecce, Bianco, et al., 2014; Valle et al., 2016).

LIMITATIONS

The present study is characterized by some limitations. First, the intervention design was limited by the use of a passive control group. In intervention research, simply belonging to an intervention group can influence outcomes due to factors such as increased attention from teachers and researchers, the novelty of the activities, and greater engagement. To address this, future studies may yield more robust conclusions by including an active control group engaged in activities that are theoretically and procedurally unrelated to ToM. This would help minimize the effects of mere exposure to an intervention and allow for a clearer attribution of any observed changes to the specific contents of the intervention.

Another limitation is the use of abbreviated versions of the ToM tasks. Indeed, school timetables allowed only brief pre- and post-test session, a common constraint in middle school. As a result, we administered shortened forms of ToM tasks that may have reduced their reliability and sensitivity to change. Future studies should allow longer assessments to employ full versions or alternative, well-validated measures that capture ToM more comprehensively.

In addition, it would be desirable to recruit a bigger sample in order to allow the comparison of the trajectories of subgroups of students, such as those who show low vs. high initial levels of ToM, also investigating gender as potential moderator. Nonetheless, due to time constraints and the ethical requirement to eventually offer the training to the passive control group, increasing the sample size of the control group was not feasible. This resulted in two groups of unequal size and contributed to the observed different variability in age and outcome measures between the groups. Although we conducted a power analysis to support our design and used statistical methods that do not assume equal variances, future studies should aim for a more balanced design to improve the reliability and generalizability of between-group comparisons.

Another limitation concerns the randomization of participants at the class level rather than individually across the two intervention groups. Although participants in our study did not differ in terms of the variables considered, geographic location, or school curriculum, class membership may have introduced biases that limit the generalizability of the findings. Future studies should aim to randomize participants at the individual level and account for potential confounding variables (e.g., teachers' experience, specific curriculum content delivered between pre- and post-test) to enhance the robustness and validity of the results.

Additionally, the present study did not include a follow-up evaluation phase to assess whether the gains were maintained over time. Therefore, future studies could introduce a third timepoint, a few months later, so that any further improvement and long-term effectiveness of the intervention could be investigated. This would also be very useful for investigating the possible far transfer effects, such as the effect of increased ToM skills on inclusion behaviors (observable and measurable during within social interactions among students at school). Additionally, social skills, which are a direct indicator of social inclusion, could also be

investigated, as they reflect an individual's ability to interact effectively with others and adapt to group social dynamics (Eisenberg & Lennon, 1983).

CONSTRAINTS TO GENERALIZABILITY

The generalizability of our findings is limited by some characteristics of the sample and study design. First, all participants were early adolescents attending public lower-secondary schools. As intervention effects can vary by developmental stage, outcomes may differ in primary, upper-secondary, or private-school settings, making replication across different age groups and school types essential. Second, all schools were located in a single region in northeastern Italy. Given the country's considerable socio-economic and cultural diversity, the present findings should be tested in other national and international contexts. Although both the intervention and outcome measures are grounded in established theoretical frameworks, successful replication will require cultural adaptation of materials and careful attention to implementation fidelity. Third, the intervention consisted of only three one-hour sessions. A higher number or longer sessions may produce different effects, and the long-term sustainability of these gains should be further tested. Finally, data collection took place shortly after the lifting of COVID-19 restrictions, when students were still affected by the pandemic's impact on their social and emotional development (Doz et al., 2023). For these reasons, future research should replicate the study across a range of geographical areas, developmental stages, intervention intensities, and post-pandemic contexts to assess the robustness and broader applicability of the results.

FUTURE DIRECTIONS

Despite its limitations, the study has strengths and innovative elements. First of all, we selected a sample of middle school students, a population that remains underexplored in the literature on interventions targeting ToM. Furthermore, the use of multiple instruments to assess ToM equally strengthens this study. While previous interventional studies have mostly used only one instrument to assess ToM (Warnell & Redcay, 2019), we decided to use more than one instrument to capture potential improvements in both socio-perceptual and socio-cognitive components of ToM. This allowed for more comprehensive and nuanced assessment of students' ToM abilities.

The study points to several directions for future perspective on ToM interventions. First of all, one relevant aspect to explore, given the digital era we live in, is how the understanding of mental states and the ability to identify relevant social information may be modulated in the context of online interactions. Furthermore, with regard to the conversational approach used during the intervention (which involved interaction between the researcher and students, and among peers), future studies could further explore the mechanisms underlying its effectiveness, as well as the potential role of moderating factors. For instance, the identity of the conversational partner (e.g., adult, peer, or group) may influence the extent to which the

conversational approach is effective in increasing ToM skills. Finally, it would be worth implementing teacher-mediated interventions focused on school diversity and inclusion, equipping educators with practical tools to effectively address students' needs and challenges, while also fostering ToM abilities and helping to move beyond stereotypical views (Valle et al., 2016).

Acknowledgments. We want to thank the schools, principal, teachers, and students for their collaboration and support. We want to thank A.B.C. Associazione per i Bambini Chirurgici del Burlo OdV ETS and Beneficentia Stiftung for the funding support received.

Author contributions. The first two listed authors contributed equally to the article, and they are first co-authors.

Conflict of interest statement. None of the authors have a conflict of interest to disclose.

Data availability statement. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

APPENDIX: EXAMPLES OF TASKS

ROLE-PLAYING ACTIVITY EXAMPLE

One of the role-playing activities proposed to students involved simulating a typical day in the life of a hospitalized child. During the activity, students were assigned specific roles and acted out a realistic scenario, recreating common interactions and dynamics within a hospital setting. In the simulation, one student played the role of the patient, while another acted as the parent, offering emotional support and accompaniment. Two additional students took on the roles of health care professionals (i.e., a doctor and a nurse) responsible for checking the patient's condition, identifying medical needs, and determining appropriate treatments. Finally, two more students played volunteers from the association, who provided recreational and playful activities to help make the patient's stay more enjoyable and engaging. This activity aimed to promote empathetic understanding of the challenges faced by hospitalized children, encouraging students to reflect on the importance of emotional support, quality medical care, and initiatives that foster the overall well-being of young patients.

EXAMPLE OF A FAUX PAS TASK

One day, Tommaso was having lunch at a restaurant when he accidentally spilled some coffee on the floor. “I’ll bring you another coffee right away,” said the waiter, and walked away. Meanwhile, Fabrizio, another customer, was standing in line at the cashier to pay the bill. Tommaso approached Fabrizio and said, “There is some coffee on the table as well, could you clean it up, please?”

Has anyone said something they shouldn’t have or something inappropriate?
☐ Yes ☐ No

If you marked Yes: answer all the following question.
If you marked No: skip to the last two questions.

Who said something they shouldn’t have or something inappropriate?

Why shouldn’t they have said it, or why was it inappropriate?

Why do you think they said it?

Does Tommaso know that Fabrizio is a customer?

What do you think Fabrizio felt?

In the story, why is Fabrizio near the cashier?

What did Tommaso spill?

EXAMPLE OF A FALSE BELIEF TASK

The pencils:

- Alberto and Benedetta are drawing with colored pencils.
- Alberto puts his pencils away in a desk drawer and goes to the kitchen to get something to drink.
- Benedetta moves the colored pencils from the desk drawer to a cupboard.
- As Alberto is returning, he sees Benedetta moving the colored pencils. Benedetta does not see Alberto.
- At this point, Alberto walks in and says he wants to get his colored pencils.

Questions:

1. Where does Benedetta think Alberto will look for the pencils?

2. Why does Benedetta think this?

3. Alberto does not want Benedetta to know that he saw her move the pencils. So where will Alberto go to retrieve his colored pencils?

4. Why?

REFERENCES

- Alloway, T. P., Bibile, V., & Lau, G. (2013). Computerized working memory training: Can it lead to gains in cognitive skills in students? *Computers in Human Behavior*, 29(3), 632–638. <https://doi.org/10.1016/j.chb.2012.10.023>
- Apperly, I. A. (2012). What is “theory of mind”? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology* (2006), 65(5), 825–839. <https://doi.org/10.1080/17470218.2012.676055>
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & 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. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 42(2), 241–251. <https://doi.org/10.1111/1469-7610.00715>
- Beauchamp, M. H., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. *Psychological Bulletin*, 136(1), 39–64. <https://doi.org/10.1037/a0017768>
- Beck, I. L. & McKeown, M. G. (2006) *Improving Comprehension with questioning the author: A fresh and expanded view of a powerful approach*. Scholastic.
- Berger, L. R., Hawks, J., Dirks, P. H., Elliott, M., & Roberts, E. M. (2017). *Homo naledi* and Pleistocene hominin evolution in subequatorial Africa. *eLife*, 6, e24234. <https://doi.org/10.7554/eLife.24234>
- Bianco, F., & Castelli, I. (2023). The promotion of mature theory of mind skills in educational settings: A mini-review. *Frontiers in Psychology*, 14, Article 1197328. <https://doi.org/10.3389/fpsyg.2023.1197328>
- Bianco, F., Lecce, S., & Banerjee, R. (2016). Conversations about mental states and theory of mind development during middle childhood: A training study. *Journal of Experimental Child Psychology*, 149, 41–61. <https://doi.org/10.1016/j.jecp.2015.11.006>
- Białycka-Pikul, M., Kołodziejczyk, A., & Bosacki, S. (2017). Advanced theory of mind in adolescence: Do age, gender and friendship style play a role? *Journal of Adolescence*, 56. <https://doi.org/10.1016/j.adolescence.2017.02.009>
- Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, 9, 267–277. <https://doi.org/10.1038/nrn2353>
- Bolte, A., & Goschke, T. (2005). On the speed of intuition: Intuitive judgments of semantic coherence under different response deadlines. *Memory & Cognition*, 33(7), 1248–1255. <https://doi.org/10.3758/BF03193226>
- Bosacki, S. L. (2000). Theory of mind and self-concept in preadolescents: Links with gender and language. *Journal of Educational Psychology*, 92(4), 709. <https://doi.org/10.1037/0022-0663.92.4.709>
- Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Frontiers in Human Neuroscience*, 6, 63. <https://doi.org/10.3389/fnhum.2012.00063>
- Caputi, M., Cugnata, F., & Brombin, C. (2021). Theory of mind and loneliness: Effects of a conversation-based training at school.

- International Journal of Psychology*, 56(2), 257–265. <https://doi.org/10.1002/ijop.12707>
- Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: the role of prosocial behavior. *Developmental Psychology*, 48(1), 257–270. <https://doi.org/10.1037/a0025402>
- Carlson, S. M., Koenig, M. A., & Harms, M. B. (2013). Theory of mind. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(4), 391–402. <https://doi.org/10.1002/wcs.1232>
- Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, 27(1), 79–151. <https://doi.org/10.1017/S0140525X04000032>
- Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science*, 14(2), F1–F10. <https://doi.org/10.1111/j.1467-7687.2010.01035.x>
- Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science*, 3(1), 81–92. <https://doi.org/10.1111/1467-7687.00102>
- Confalonieri, F. (2009). La Teoria della Mente nell'adulto: applicazione di prove complesse di falsa credenza [Theory of mind in adults: Assessing false-belief reasoning]. *Imparare*, 6.
- Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews. Neuroscience*, 13(9), 636–650. <https://doi.org/10.1038/nrn3313>
- Cuder, A., Vidoz, M., De Vita, C., Pellizzoni, S., & Passolunghi, M. C. (2022). Numerical training videos and early numerical achievement: A study on 3-year-old preschoolers. *Brain Sciences*, 12(1), 88. <https://doi.org/10.3390/brainsci12010088>
- Cutting, A. L., & Dunn, J. (2002). The cost of understanding other people: Social cognition predicts young children's sensitivity to criticism. *Journal of Child Psychology and Psychiatry*, 43(7), 849–860. <https://doi.org/10.1111/1469-7610.t01-1-00047>
- Dehning, S., Girma, E., Gasperi, S., Meyer, S., Tesfaye, M., & Siebeck, M. (2012). Comparative cross-sectional study of empathy among first year and final year medical students in Jimma University, Ethiopia: Steady state of the heart and opening of the eyes. *BMC Medical Education*, 12, 34. <https://doi.org/10.1186/1472-6920-12-34>
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology*, 30(1), 92–101. <https://doi.org/10.5334/irsp.82>
- Delacre, M., Lakens, D., Ley, C., Liu, L., & Leys, C. (2021). Why Hedges' g's based on the non-pooled standard deviation should be reported with Welch's t-test. *PsyArXiv*. <https://doi.org/10.31234/osf.io/tu6mp>
- Denham, S. A. (2006). Social-emotional competence as support for school readiness: What is it and how do we assess it? *Early Education and Development*, 17(1), 57–89. https://doi.org/10.1207/s15566935eed1701_4
- Deolmi, M., & Pisani F. (2020). Psychological and psychiatric impact of COVID-19 pandemic among children and adolescents. *Acta Biomedica*, 91(4), Article e2020149. <https://doi.org/10.23750/abm.v91i4.10870>
- Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: ToM, gender, and social experiences in middle childhood. *Child Development*, 84(3), 989–1003. <https://doi.org/10.1111/cdev.12017>
- Diaz, V., & Farrar, M. J. (2018). The missing explanation of the false-belief advantage in bilingual children: A longitudinal study. *Developmental Science*, 21(4), e12594. <https://doi.org/10.1111/desc.12594>
- Doz, E., Cuder, A., Caputi, M., Pellizzoni, S., & Passolunghi, M. C. (2023). Distance learning environment: Perspective of Italian primary and secondary teachers during COVID-19 pandemic. *Learning Environments Research*, 26(2), 555–571. <https://doi.org/10.1007/s10984-022-09451-9>
- Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind

- continues to develop in late adolescence. *Developmental Science*, 13(2), 331–338. <https://doi.org/10.1111/j.1467-7687.2009.00888.x>
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, 94(1), 100–131. <https://doi.org/10.1037/0033-2909.94.1.100>.
- Eisenmajer, R., & Prior, M. (1991). Cognitive linguistic correlates of 'theory of mind' ability in autistic children. *British Journal of Developmental Psychology*, 9(2), 351–364. <https://doi.org/10.1111/j.2044-835X.1991.tb00882.x>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Filippello, P., Sorrenti, L., and Gugliandolo, M. C. (2022). Teoria della mente e problem solving sociale [Theory of mind and social problem solving]. In I. Castelli, A. Marchetti, & D. Massaro (Eds.), *La competenza sociale nei bambini: Valutare e promuovere la teoria della mente in età prescolare e scolare* [Social competence in children: Evaluating and enhancing theory of mind in preschool and school age] (pp. 337–369). Junior-Bambini.
- Fiske, S. T. (2009). From dehumanization and objectification, to rehumanization: Neuroimaging studies on the building blocks of empathy. *Annals of the New York Academy of Sciences*, 1167, 31–34. <https://doi.org/10.1111/j.1749-6632.2009.04544.x>
- Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, 78(4), 708–724. <https://doi.org/10.1037/0022-3514.78.4.708>
- Gao, Q., Huang, Q., Zhang, Q., & Wei, C. (2020). Does executive function influence the development of theory of mind in elementary students? *Current Psychology*, 39(2), 389–396. <https://doi.org/10.1007/s12144-018-0107-1>
- Greenberg, A., Bellana, B., & Bialystok, E. (2013). Perspective-taking ability in bilingual children: Extending advantages in executive control to spatial reasoning. *Cognitive Development*, 28(1), 41–50. <https://doi.org/10.1016/j.cogdev.2012.10.002>
- Greenfield, P. M. (2000). Social and emotional development in childhood. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (5th ed., Vol. 3, pp. 799–878). John Wiley.
- Gregory, C., Lough, S., Stone, V.E., Erzincioğlu, S., Martin, L., Baron-Cohen, S. & Hodges, J. (2002). Theory of mind in frontotemporal dementia and Alzheimer's disease: Theoretical and practical implications. *Brain*, 125, 752–764. <https://doi.org/10.1093/brain/awf079>
- Grice, P. (1989). *Studies in the way of words*. Harvard University Press.
- Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24(2), 129–154. <https://doi.org/10.1007/BF02172093>
- Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind & Language*, 7(1–2), 120–144. <https://doi.org/10.1111/j.1468-0017.1992.tb00201.x>
- Higgins, W. C., Kaplan, D. M., Deschrijver, E., & Ross, R. M. (2024). Construct validity evidence reporting practices for the Reading the Mind in the Eyes Test: A systematic scoping review. *Clinical Psychology Review*, 108, 102378. <https://doi.org/10.1016/j.cpr.2023.102378>
- Higgins, W. C., Ross, R. M., Langdon, R., & Polito, V. (2022). The "Reading the Mind in the Eyes" Test shows poor psychometric properties in a large, demographically representative U.S. sample. *Assessment*, 30(6), 1777–1789. <https://doi.org/10.1177/10731911221124342>
- Hoffman, M. L. (2001). Toward a comprehensive empathy-based theory of prosocial moral development. In A. C. Bohart & D. J. Stipek (Eds.), *Constructive & destructive behavior: Implications for family, school, & society* (pp. 61–86). American Psychological Association. <https://doi.org/10.1037/10433-003>
- Hofmann, S. G., Doan, S. N., Sprung, M., Wilson, A., Ebesutani, C., Andrews, L. A.,

- Curtiss, J., & Harris, P. L. (2016). Training children's theory-of-mind: A meta-analysis of controlled studies. *Cognition*, 150, 200–212. <https://doi.org/10.1016/j.cognition.2016.01.006>
- Honoré, N., Houssa, M., Volckaert, A., Noël, M. P., & Nader-Grosbois, N. (2020). Training inhibition and social cognition in the classrooms. *Frontiers in Psychology*, 11, 1974. <https://doi.org/10.3389/fpsyg.2020.01974>
- Hoyo, Á., Rueda, M. R., & Rodríguez-Bailón, R. (2019). Children's individual differences in executive function and theory of mind in relation to prejudice toward social minorities. *Frontiers in Psychology*, 10, 2293. <https://doi.org/10.3389/fpsyg.2019.02293>
- Hughes, C. (2011). *Social understanding and social lives: From toddlerhood through to the transition to school*. Psychology Press.
- Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, 13(4), 590–619. <https://doi.org/10.1111/j.1467-9507.2004.00285.x>
- Jha, M. & Sharma, V. (2020). Promoting theory of mind: Intervention and training. *Delhi Psychiatry Journal*, 23(2), 7–10.
- Kittel, A. F. D., Olderbak, S., & Wilhelm, O. (2021). Sty in the mind's eye: A meta-analytic investigation of the nomological network and internal consistency of the "Reading the Mind in the Eyes" Test. *Assessment*, 29(5), 872–895. <https://doi.org/10.1177/1073191121996469>
- Lecce, S., Bianco, F., Devine, R. T., Hughes, C., & Banerjee, R. (2014). Promoting theory of mind during middle childhood: A training program. *Journal of Experimental Child Psychology*, 126, 52–67. <https://doi.org/10.1016/j.jecp.2014.03.002>
- Lecce, S., Caputi, M. & Pagnin, A. (2014). False-belief understanding at age 5 predicts beliefs about learning in year 3 of primary school. *European Journal of Developmental Psychology*, 12(1), 40–53. <https://doi.org/10.1080/17405629.2014.949665>
- Leekam, S. R., & Prior, M. (1994). Can autistic children distinguish lies from jokes? A second look at second-order belief attribution. *Journal of Child Psychology and Psychiatry*, 35(5), 901–915. <https://doi.org/10.1111/j.1469-7610.1994.tb02301.x>
- Li, X., Wang, K., Wang, F., Tao, Q., Xie, Y., & Cheng, Q. (2013). Aging of theory of mind: The influence of educational level and cognitive processing. *International Journal of Psychology*, 48(4), 715–727. <https://doi.org/10.1080/00207594.2012.673724>
- Liverta-Sempio, O., Marchetti, A., & Lecciso, F. (2005). *Faux pas: traduzione italiana* [Faux pas: Italian tradition]. Theory of Mind Research Unit, Department of Psychology, Catholic University of the Sacred Heart, Milan.
- Lombardi, E., Valle, A., Bianco, F., Castelli, I., Massaro, D., & 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. *Cognition and Emotion*, 36(5), 975–986. <https://doi.org/10.1080/02699931.2022.2067521>
- Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, 62, 103–134. <https://doi.org/10.1146/annurev-psych-120709-145406>
- Meinhardt-Injac, B., Daum, M. M., Meinhardt, G., & Persike, M. (2018). The two-systems account of theory of mind: Testing the links to social-perceptual and cognitive abilities. *Frontiers in Human Neuroscience*, 12(25). <https://doi.org/10.3389/fnhum.2018.00025>
- Meinhardt-Injac, B., Daum, M. M., & Meinhardt, G. (2020). Theory of mind development from adolescence to adulthood: Testing the two-component model. *British Journal of Developmental Psychology*, 38(2), 289–303. <https://doi.org/10.1111/bjdp.12320>
- Melot, A., & Angeard, N. (2003). Theory of mind: Is training contagious. *Developmental Science*, 6(2), 178–184. <https://doi.org/10.1111/1467-7687.00269>
- Navarro, E., & Conway, A. R. (2021). Adult bilinguals outperform monolinguals in theory of mind. *Quarterly Journal of Experimental Psychology*, 74(11), 1841–1851. <https://doi.org/10.1177/17470218211009159>
- Navarro, E., Goring, S. A., & Conway, A. R. A. (2021). The relationship between

- theory of mind and intelligence: A formative g approach. *Journal of Intelligence*, 9(1), 11. <https://doi.org/10.3390/jintelligence9010011>
- Nelson, K. (2005). Language pathways into the community of minds. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 26–49). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195159912.003.0002>
- Nota, L., Ginevra, M. C., & Santilli, S. (2015). Life design and prevention. In L. Nota, & J. Rossier (Eds.), *Handbook of life design: From practice to theory and from theory to practice* (pp. 183–199). Hogrefe.
- Ornaghi, V., Brockmeier, J., & Grazzani, I. (2014). Enhancing social cognition by training children in emotion understanding: A primary school study. *Journal of Experimental Child Psychology*, 119, 26–39. <https://doi.org/10.1016/j.jecp.2013.10.005>
- Quesque, F., Apperly, I., Baillargeon, R., Baron-Cohen, S., Becchio, C., Bekkering, H., Bernstein, D., Bertoux, M., Bird, G., Bukowski, H., Burgmer, P., Carruthers, P., Catmur, C., Dziobek, I., Epley, N., Erle, T. M., Frith, C., Frith, U., Galang, C. M., . . . Brass, M. (2024). Defining key concepts for mental state attribution. *Communications Psychology*, 2(1), 1–5. <https://doi.org/10.1038/s44271-024-00077-6>
- Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nature Reviews Psychology*, 1(4), 223–235. <https://doi.org/10.1038/s44159-022-00037-z>
- Romeo, R. D. (2013). The teenage brain: The Stress response and the adolescent brain. *Current Directions in Psychological Science*, 22(2), 140–145. <https://doi.org/10.1177/0963721413475445>
- Ronchi, L., Banerjee, R., & Lecce, S. (2020). Theory of mind and peer relationships: The role of social anxiety. *Social Development*, 29, 478–493. <https://doi.org/10.1111/sode.12417>
- Rubio-Fernández, P., & Glucksberg, S. (2012). Reasoning about other people's beliefs: Bilinguals have an advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(1), 211–217. <https://doi.org/10.1037/a0025162>
- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. *Behavioral Ecology*, 17(4), 688–690. <https://doi.org/10.1093/beheco/ark016>
- Sabato, H., Abraham, Y. E. Kogut, T. (2021). Too lonely to help: Early adolescents' social connections and willingness to help during Covid-19 lockdown. *Journal of Research and Adolescence*, 31, 764–779. <https://doi.org/10.1111/jora.12655>
- Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, 19(2), 65–72. <https://doi.org/10.1016/j.tics.2014.11.007>
- Serafin, M., & Surian, L. (2004). Il test degli occhi: uno strumento per valutare la "Teoria della Mente" [The Eyes Test: An instrument to assess the "Theory of Mind"]. *Giornale Italiano di Psicologia*, 31(4), 839–862. <https://doi.org/10.1421/18849>
- Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, 46(5), 668–677. <https://doi.org/10.1016/j.cortex.2009.04.008>
- Sodian, B. (1991). The development of deception in young children. *British Journal of Developmental Psychology*, 9(1), 173–188. <https://doi.org/10.1111/j.2044-835X.1991.tb00869.x>
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell.
- Stein, H. (2006). Does mentalizing promote resilience? In J. Allen & P. Fonagy (Eds.), *Handbook of mentalization-based treatment* (pp. 307–326). John Wiley.
- Stone, V. E., Baron-Cohen, S., Calder, A., Keane, J., & Young, A. (2003). Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia*, 41(2), 209–220. [https://doi.org/10.1016/S0028-3932\(02\)00151-3](https://doi.org/10.1016/S0028-3932(02)00151-3)
- Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, 10, 640–656. <https://doi.org/10.1162/089892998562942>
- Symons, D. K. (2004). Mental state discourse, theory of mind, and the internalization of self-other understanding. *Developmental*

- Review*, 24(2), 159–188. <https://doi.org/10.1016/j.dr.2004.03.001>
- Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: Evidence from Williams syndrome. *Cognition*, 76(1), 59–90. [https://doi.org/10.1016/S0010-0277\(00\)00069-X](https://doi.org/10.1016/S0010-0277(00)00069-X)
- Thurstone, T. G. (1962). *Primary mental abilities*. Science Research Associates.
- Turnbull, W., & Carpendale, J. I. (1999). A social pragmatic model of talk: Implications for research on the development of children's social understanding. *Human Development*, 42(6), 328–355. <https://doi.org/10.1159/000022641>
- Valle, A., Massaro, D., Castelli, I., Intra, F. S., Lombardi, E., Bracaglia, E., & Marchetti, A. (2016). Promoting mentalizing in pupils by acting on teachers: Preliminary Italian evidence of the “Thought in Mind” Project. *Frontiers in Psychology*, 7, Article 1213. <https://doi.org/10.3389/fpsyg.2016.01213>
- Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Vezzali, L., Stathi, S., & Giovannini, D. (2012). Indirect contact through book reading: Improving adolescents' attitudes and behavioral intentions toward immigrants. *Psychology in the Schools*, 49(2), 148–162. <https://doi.org/10.1002/pits.20621>
- Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, 191(5), Article 103997. <https://doi.org/10.1016/j.cognition.2019.06.009>
- Weimer, A. A., Parault Dowds, S. J., Fabricius, W. V., Schwanenflugel, P. J., & Suh, G. W. (2017). Development of constructivist theory of mind from middle childhood to early adulthood and its relation to social cognition and behavior. *Journal of Experimental Child Psychology*, 154, 28–45. <https://doi.org/10.1016/j.jecp.2016.10.002>
- Wellman, H. M., & Peterson, C. C. (2013). Deafness, thought bubbles, and theory-of-mind development. *Developmental Psychology*, 49(12), 2357–2367. <https://doi.org/10.1037/a0032419>
- Wood, D., & Middleton, D. (1975). A study of assisted problem-solving. *British Journal of Psychology*, 66(2), 181–191. <https://doi.org/10.1111/j.2044-8295.1975.tb01454.x>
- Yeung, E. K. L., Apperly, I. A., & Devine, R. T. (2024). Measures of individual differences in adult theory of mind: A systematic review. *Neuroscience and Biobehavioral Reviews*, 157, 105481. <https://doi.org/10.1016/j.neubiorev.2023.105481>