

The Effect of Peers' Theory of Mind on Children's Own Theory of Mind development

Lecce, Serena; Ronchi, Luca; Devine, R.T.

DOI:

[10.1037/dev0001758](https://doi.org/10.1037/dev0001758)

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Lecce, S, Ronchi, L & Devine, RT 2024, 'The Effect of Peers' Theory of Mind on Children's Own Theory of Mind development: A Longitudinal Study in Middle Childhood and Early Adolescence', *Developmental Psychology*.
<https://doi.org/10.1037/dev0001758>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

©American Psychological Association, 2024. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. The final article is available, upon publication, at: <https://psycnet.apa.org/doi/10.1037/dev0001758>

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

**The Effect of Peers' Theory of Mind on Children's Own Theory of Mind development:
A Longitudinal Study in Middle Childhood and Early Adolescence**

Serena Lecce¹, Luca Ronchi¹, and Rory T. Devine²

¹ Department of Brain and Behavioural Sciences, University of Pavia, Italy

² Centre for Developmental Science, School of Psychology, University of Birmingham, UK

Author accepted manuscript to appear in *Developmental Psychology*.

Manuscript accepted on 26th February 2024.

Author Note

We have no known conflict of interest to disclose. Correspondence concerning this article should be addressed to Serena Lecce, Department of Brain and Behavioural Science, University of Pavia, P.zza Botta 6, 27100 Pavia, Italy. Email: slecce@unipv.it

Abstract

While there is considerable evidence that children's early ability to understand others' mental states, called 'theory of mind', is shaped by family experiences, it remains unclear whether children's social interactions at school influence theory of mind beyond early childhood. We tested whether the mean level ('quantity') and/or the diversity ('variety') of peers' theory of mind influenced children's own theory of mind. We also examined whether peer effects on theory of mind were independent of possible confounding variables (e.g., verbal ability, social isolation) and comparable across children with different initial levels of theory of mind and social status. 454 8- to 12-year-old children completed assessments of theory of mind and peer and friendship nominations at baseline and (for theory of mind only) 1 year later. The variety (but not the quantity) of peers' theory of mind predicted the development of children's theory of mind over and above control variables. The magnitude of the peer effect was comparable across different levels of children's theory of mind and between children indexed as socially isolated and those who were not. These findings fit with socio-cultural models and highlight the importance of school environment in the development of theory of mind.

Keywords: peers, theory of mind, diversity, social isolation

Public Statement: This study delves into the critical question of whether peers influence the development of 'theory of mind' defined as children's ability to understand others' mental states such as desires, beliefs and emotions. Involving 454 children aged 8 to 12, the research reveals that belong to a classroom with greater heterogeneity in theory of mind performance predicts the advancement of children's own theory of mind, irrespective of other factors such as verbal ability and social isolation. These findings underscore the importance of the school environment in fostering the development of fundamental social cognitive skills, with significant implications for enhancing children's educational experiences.

**The Effect of Peers' Theory of Mind on Children's Own Theory of Mind development:
A Longitudinal Study in Middle Childhood and Early Adolescence**

The past decade has seen an expansion of research on theory of mind (i.e., the ability to reason about others' thoughts, feelings, and desires) beyond the traditional confines of early childhood into middle childhood (Devine & Lecce, 2021). This shift in focus has prompted consideration about links between individual differences in theory of mind and children's life at school (Devine & Lecce, 2021). The over-arching aim of the current study was to examine, for the first time, the effect of peers' theory of mind on children's own theory of mind using a longitudinal design. Understanding whether and how peers, and more generally school environments, influence the development of children's theory of mind is vital because children who excel at reasoning about others' minds during the primary school years are more likely to be accepted by their peers (Slaughter, Imuta, Peterson, & Henry, 2015), to be viewed as socially skilled by teachers (Devine & Apperly, 2022), and show higher academic achievement (Lecce, 2021), especially in the area of reading comprehension (Lecce, Bianco, & Hughes, 2021) and scientific reasoning (Osterhaus & Koerber, 2023).

Starting from the work of Wundt (1900/1921) who argued that "folk psychology" emerges from group experiences, numerous 'social' accounts on theory of mind development have emerged in the literature (Hughes & Devine, 2015). These social constructivist theories place children's social interactions with other members of their culture at the heart of mindreading development and view social relationships as the path through which children construct an understanding of others' minds (Carpendale & Lewis, 2006). While these theories are supported by evidence showing the effects of early social experiences within the family on young children's emerging theory of mind (Devine & Hughes, 2018), there is a scarcity of research on how school environments impact on children's ongoing theory of mind development. This is problematic for two reasons. First, by focusing on parent-child interactions, it is difficult to disentangle the relative influence of

environmental factors and genetics on theory of mind. Second, by emphasising early childhood, it is unclear how social experiences affect the ongoing development of theory of mind.

The current study addresses this gap in the literature by examining the influence of peers' theory of mind on children's own theory of mind. While numerous studies have investigated how classroom peers' language, self-regulation, and academic skills influence children's own abilities (Mashburn, Justice, Downer, & Pianta, 2009; Yang et al., 2023), to date, researchers have yet to investigate how peers' theory of mind shapes children's own theory of mind. There are at least two ways that peers' theory of mind might influence children's theory of mind. First, children's theory of mind might be influenced by the average level of peers' theory of mind. According to this 'quantity' hypothesis, children in classrooms with high average levels of theory of mind performance will show a greater increase in theory of mind over time than children with peers who have low average levels of theory of mind. If, for example, there is a high mean level of ToM in the classroom, this could reflect a greater quantity of exposure to ToM. Second, children belonging to classrooms with greater diversity or heterogeneity in theory of mind performance might develop more advanced theory of mind than children belonging to classrooms with less diversity (Yang et al., 2023). According to this 'variety' hypothesis, greater diversity in peers' theory of mind will provide opportunities for children to experience differences in perspectives, driving greater gains in theory of mind.

Theoretical Perspectives on Peer Effects

One core aim of the present study was to test the quantity hypothesis. Although no research has directly tested how average levels of classroom peers' theory of mind influence children's own theory of mind, indirect support comes from at least three strands of evidence. First, studies of conversations have shown that the overall amount of mental state language a child is exposed to within the family predicts children's ability to infer others' mental states over time (Hughes & Devine, 2019). The effect of frequency of mental state conversations extends to middle childhood. Lecce et al. (2021) have shown that primary school teachers' propensity for mental-state language

was uniquely associated with pupils' theory of mind even when models were adjusted for child-related (i.e., age, verbal ability, number of siblings and SES) and teacher-related variables (i.e., theory of mind, verbal ability and years of experience).

Second, theory of mind training studies bolster the quantity hypothesis. Increasing children's exposure to mental state reasoning via dedicated training programs has a positive effect on children's theory of mind performance (Hofmann et al., 2016). Among school-aged children, interventions prompting children to reason about mental states via narratives and involving them in group conversations about the characters' mental states increased children's theory of mind both in the short (Lecce, Bianco, Devine, Hughes, & Banerjee, 2014) and medium term (Bianco & Lecce, 2016).

Third, support for the quantity hypothesis also comes from research on siblings. For example, Jenkins, Turrell, Kogushi, Louis, and Ross (2003) found that children with an older sibling were exposed to a greater quantity of mental state talk than children with a younger sibling. The positive effect of having an older sibling on theory of mind has also been found in middle childhood. For example, over and above the effects of child age and executive function, the number of older siblings predicted performance on theory of mind tasks in 4- to 11-year-old children (Kennedy, Lagattuta, & Sayfan, 2015). These findings suggest that daily interactions with a more knowledgeable partner expose younger children to a greater amount of mental state conversations, providing an ideal training ground for theory-of-mind development.

According to the variety hypothesis, theory-of-mind development is more likely to occur among children whose peers show greater variation in theory of mind. This view is grounded in the constructivist concept of cognitive disequilibrium (Piaget, 1977a), whereby exposure to conflicting events or sources of information stimulates intellectual development. Translating this mechanism into the domain of theory of mind, children who are exposed to a greater variety of perspectives about the same reality should have more frequent opportunities to reflect on the subjective nature of mental states and on the link between these subjective experiences and social behavior. So,

interactions with peers with varying levels of mindreading ability might stimulate the refinement of theory of mind. Although this hypothesis has not been directly tested, there is indirect evidence to support this position. In a study involving children aged 3 to 5 years, the presence of multiple child siblings of varied ages (and presumably, varying levels of theory mind) predicted theory of mind performance (Peterson, 2000). Furthermore, children with a sibling performed better at false belief tasks than children with a twin (Cassidy, Fineberg, Brown, & Perkins, 2005) and monozygotic twins performed worse than dizygotic twins (Deneault et al., 2008). Overall, these data point to the importance of interacting with siblings who vary in age or ability for theory of mind.

Findings from at least three other studies fit with the variety hypothesis. First, children belonging to mixed-aged classrooms in kindergarten had better false-belief understanding than those attending same-age only classrooms (Wang & Su, 2009). Second, typically developing children (aged 7- to 9-years old) attending inclusive classrooms (made up of children with and without disabilities) exhibited greater gains in theory of mind than children in non-inclusive classrooms over two years (Smogorzewska, Szumski, & Grygiel, 2020). Third, in a recent study, 8- to 13-year-old children with cross-ethnic friendships showed better performance on theory of mind measures than children without cross-ethnic friendships. These findings imply that classroom diversity can foster children's theory of mind (Devine, Traynor, Ronchi, & Lecce, 2024).

In testing the amount and the variety hypotheses we were also interested in examining whether any effects were contingent on children's initial level of theory of mind and their relationships with peers in the classroom. Existing research shows that peer influences on children's cognitive development can vary according to children's own skills (Burke & Sass, 2013) with children who have low ability often showing greater benefits (Hanushek et al., 2003). For example, children with limited language skills seemed particularly to benefit from being in classrooms with peers who had higher levels (Justice, Petscher, Schatschneider, & Mashburn, 2011) or greater variability in English (Aikens et al., 2010). The view that initially low performing children take advantage of their social environment fits with socio-cultural accounts, which claim that

development is driven by interactions with more knowledgeable others (Vygotsky, 1978). With regard to theory of mind, Jenkins and Astington (1996) found that the positive effect of the presence of siblings on children's false-belief performance was restricted to verbally less able children, suggesting that the presence of siblings can compensate for slower language development in developing false belief understanding. These results highlight an interaction between social experience and children's own cognitive characteristics. The effect of peers' theory of mind may also depend on children's peer relationships. It is plausible that peers' theory of mind affects children's own theory of mind only if children can actively participate in those peer interactions. This view fits with data showing that reduced opportunities to engage in peer relationships hinder the acquisition of mental state reasoning (Banerjee, Watling, & Caputi, 2011). We addressed this issue by examining whether social isolation, characterized as the lack of mutual nominations for being 'most liked' or having friendships within the classroom, moderates the effect of peers' theory of mind.

The present study

While previous literature has examined the links between theory of mind and peer relationships (Slaughter et al., 2015), the present study breaks new ground by testing the existence and specificity of the influence of peers' theory of mind on children's own theory of mind. We investigated whether the average level of theory of mind in a classroom (i.e., quantity hypothesis) and/or the diversity of theory of mind ability in a classroom (i.e., variety hypothesis) predicted individual differences in children's theory of mind one year later. We controlled for rank-order stability in theory of mind over time and a number of covariates known to be related to theory of mind performance (i.e., age, gender, SES, verbal ability, social isolation). In evaluating the specificity of any effects of peers' theory of mind we also tested whether the peer effect varied as a function of children's initial theory of mind and social isolation. Finally, to address the possibility that peers' theory of mind might reflect domain-general contextual effects on children's theory

mind (rather than effects of theory of mind *per se*), we examined whether peers' theory of mind was related to children's verbal ability.

Method

Participants

We recruited 454 Italian children between the ages of 8 and 12 years from 20 classrooms in state-funded primary schools across Northern Italy. Of these 454 children, 45 were excluded because their caregivers did not provide consent for their participation and/or the children were unable to participate in the study unaided by a classroom assistant. The proportion of participating children from each classroom in the final sample was 89.42% on average ($SD = 10.21\%$, range 68.4 – 100%). The 409 (49.4% girls) children in the final sample were aged between 7.95 and 12.20 (M age = 9.43, $SD = 0.68$) at the first time point of the data collection. According to the Italian school system, 72 children (3 classrooms, M age = 8.55, $SD = 0.31$) were enrolled in Year 3 of primary school, 264 (14 classrooms, M age = 9.34, $SD = 0.31$) were enrolled in Year 4, and 73 children (3 classrooms, M age = 10.51, $SD = 0.40$) were enrolled in Year 5 at the first wave of data collection (Time 1). Except for 19 children belonging to the same Year 4 classroom, whose main teacher withdrew at Time 2, all remaining children participated again approximately 1 year later (Time 2). Regarding children's ethnicity, 94% identified as White, 1% as Asian, 0.5% as Black, 2.9% as Mixed Race, and 1.7% as Other. Nine per cent of participating Italian children had a formal statement of special educational needs and disability. Classroom size ranged from 15 to 28 pupils ($M = 23.11$, $SD = 2.98$).

Procedure

Parental consent was obtained prior to data collection. Data were part of two larger longitudinal studies examining social and cognitive correlates of children's theory of mind in middle childhood. This paper presents data collected during the first wave of data collection (November 2015 for Study 1 and November 2017 for Study 2) and one year later. The mean test-retest gap between time points was 325.5 days ($SD = 33.6$). Children were tested at school during

two half-hour sessions, 1 week apart. At each time point children completed two different theory of mind tasks, the Strange Stories (White, Hill, Happé, & Frith, 2009) and Silent Films tasks (Devine & Hughes, 2013) and a friendships assessment questionnaire (Parker & Asher, 1993). Children also completed a verbal ability test (Primary Mental Abilities, PMA; Thurstone & Thurstone, 1962) and a socioeconomic status questionnaire (Family Affluence Scale; FAS; Currie et al., 2008) at Time 1. One researcher introduced the tasks and read the questions out loud one at a time, supported by a PowerPoint presentation. A second researcher in the classroom ensured that all participants understood the task instructions and completed their response booklets individually. Ethical approval for this study was granted by the Ethical committee of the University of Pavia (N° 031/2019).

Measures

Theory of Mind. We assessed children's theory of mind using two age-appropriate tasks: the Silent Film Task (Devine & Hughes, 2013) and the Strange Stories Task (Happé, 1994). In the *Silent Film Task* (Devine & Hughes, 2013) children watched five short film clips from a classic silent comedy depicting instances of deception, misunderstanding, and false belief. Children responded to a single question about each clip (read aloud by the research assistant), which required an explanation of a character's behavior. The research assistant did not play the next clip until all children had recorded an answer. Children's open-ended responses were later scored by two trained research assistants. Children received 2 points for accurate mentalizing given the context, 1 point for partially correct responses, and 0 points for inaccurate or irrelevant responses (Devine et al., 2023). A second rater independently coded 25% of the responses at each time point, and interrater agreement was established using Cohen's kappa (at T1, $k = .81$; at T2, $k = .81$). In the *Strange Stories Task* (Happé, 1994), the researcher read aloud four short vignettes, involving deception (1 story), misunderstanding (2 stories) or double bluff (1 story). The stories were displayed on a large screen for the children to see. Children answered an open-ended question about the characters' behavior. The researcher only showed the next story when all children had recorded their response.

Two trained research assistants later coded these written responses. Correct responses involving accurate mentalizing received 2 points, partially correct responses received 1 point, and inaccurate or irrelevant responses scored 0 points (see White et al., 2009 for details on coding). A second coder independently scored 25% of the responses at each time point. The interrater agreement was established using Cohen's kappa (at T1, $k = .91$; at T2, $k = .88$).

The Silent Film and the Strange Stories tasks have been used in large-scale studies of children's theory of mind across middle childhood (e.g., Lecce, Ronchi, & Devine, 2022). They show strong convergent validity, over and above variation in children's verbal ability, SES, or narrative comprehension, and load onto a single theory-of-mind latent factor (Devine & Hughes, 2016). Based on this literature and on Confirmatory Factor Analyses conducted on the present data (see the results section) we created a composite theory of mind score by combining children's scores on the Silent Film and the Strange Stories tasks. These theory-of-mind tasks were used to assess individuals' theory of mind levels and also to derive a measure of classroom theory of mind. To index classroom theory of mind, we considered a measure of central tendency and variation. Given that the number of children belonging to the same classroom was under 30, we selected the median and interquartile range as these indices were less sensitive to outliers than the mean and the standard deviation.

Social Isolation. To measure social isolation we combined the sociometric positive peer nomination procedure (Coie, Dodge, & Coppotelli, 1982) with the mutual friendship nomination procedure described by Parker and Asher (1993). Children were asked to nominate their top three 'friends' in the class and up to three classmates they 'most like spending time with'. Children were considered as being socially isolated if they had zero reciprocated nominations both in the friendship and the positive peer nominations ($N = 42$; 10%). With the exception of 9 children, whose parents did not give consent to participate, all other children (including those who were not eligible due to learning disabilities or not being native Italian speakers, see the participants section) took part in the nomination procedure as both nominators and nominees. This procedure allowed us

to maximize the number of nominators to collect reliable and valid data (Cillessen & Marks, 2017), and to have a comprehensive picture of the social relations in the classroom.

Socioeconomic status. The Family Affluence Scale (FAS - Currie et al., 2008) is a short self-report measure of material wealth developed as part of the WHO Health Behaviour in School-Aged Children Study to investigate the SES in youth (Currie et al., 2008). Previous studies indicated high agreement between 11-year-old children's and their parents' reports on the FAS (Andersen et al., 2008). There are four questions about the following: family car ownership (range: 0–2), the participants having/not having their own unshared bedroom (range: 0–1), the number of computers at home (range: 0–3), and the number of times participants went on a vacation during the past year (range: 0–3). Responses were summed into an index of family affluence (possible range: 0–9). The Cronbach's α was .40, a value that is in line with existing studies (Schnohr et al., 2008) and likely due to the limited number of items (Torsheim et al., 2016).

Verbal Ability. Children's verbal ability was measured using two receptive vocabulary tasks: the *Mill Hill Vocabulary Scale* (Rust, 2008) for Study 1 and the *Primary Mental Abilities Vocabulary Scale* (Rubini & Rossi, 1982; Thurstone & Thurstone, 1962) for Study 2. The *Mill Hill Vocabulary Scale* requires children to select a synonym for 20 target words from six possible response options each and received 1 point for each correctly identified word. The *Primary Mental Abilities Vocabulary Scale* requires children to select a synonym for 50 target words from five possible response options and received 1 point for each correctly identified word. To control for the difference in the number of items and in the degree of difficulty between the different vocabulary tasks used, children's total scores were age residualized and standardized within the vocabulary test used. The age-residualized and standardized verbal ability scores were interpreted as the child's score deviation (positive or negative) from the average score of children of the same age, who completed the same vocabulary test. Reliability coefficients measured using the Kuder-Richardson Formula 20 (Kuder & Richardson, 1937) were .93 for the *Mill Hill Vocabulary Scale* and .80 for the *Primary Mental Abilities Vocabulary Scale*.

Analysis Plan

We conducted Confirmatory Factor Analyses (CFA) using Mplus version 7 (Muthén & Muthén, 2017) to test for longitudinal measurement invariance of the theory of mind battery and to evaluate mean level change in theory of mind. Establishing measurement invariance for the theory of mind latent factor over time is a prerequisite to estimating differences in latent means over time (Brown, 2015). We tested measurement invariance by imposing equality restraints to the measurement model across time and evaluating the change in model fit at subsequent steps (Brown, 2015). Given the categorical nature of indicators, we used a mean- and variance-adjusted weighted least squares (WLSMV) estimator with Delta parameterisation (Muthén & Muthén, 2017). We evaluated model fit using three primary criteria: Comparative Fit Index (CFI) > .90, Tucker Lewis Index (TLI) > .90, Root Mean Square Error of Approximation (RMSEA) < .08 (Brown, 2015). Nested model comparisons were deemed as nonsignificant using the following criteria: $\Delta\text{CFI} < .010$ and $\Delta\text{RMSEA} < .010$ (Cheung & Gardner, 2015).

We used Mixed Linear Modelling (MLM) to test whether classroom-related contextual factors (i.e., the mean level of peers' theory of mind and/or the variety of peers' theory of mind) uniquely predicted variation in children's theory of mind one year later. Given that children were nested within classrooms, we specified a two-level hierarchical structure for our regression equation in which children represented the lower level of analysis (level 1) and classrooms represented the upper-level clustering variable (level 2) to account for nonindependence of the observations. First, we ran a *Random Intercepts Empty model* (M0) in which we included the outcome variable (i.e., children's theory of mind at Time 2) and allowed the intercept (i.e., theory of mind mean) to vary across classrooms. This model allowed us to partition the variance of the outcome variable into the within- and between-classroom components and estimate the percentage of total variance in theory of mind scores attributable to belonging to a particular classroom (i.e., Intraclass Correlation Coefficient – ICC). The ICC quantifies the magnitude of contextual influence of the classroom environment on children's theory of mind (Merlo, Wagner, Austin, Subramanian, & Leckie, 2018).

We then took a three-step approach to address our main question. First, we estimated a Random Intercept model (M1), in which we included child-level theory of mind scores at Time 1 together with covariates as predictors of Time 2 theory of mind. In this model mean-level ToM scores across classrooms were adjusted for variation in children's age, gender, SES, verbal ability, social isolation, and presence of special education needs. Second, we added the classroom theory of mind median and IQR as level 2 predictors. This second Random Intercept model (M2) tested the unique effects of peers' theory of mind on children's theory of mind at Time 2 over and above cross-time stability in children's theory of mind. We then tested two models in which we added interaction terms between classroom level variables (theory of mind median and IQR) and children's theory of mind at time 1 (M3a) or children's social isolation (M3b). This allowed us to test if any contextual effect of peers' theory of mind on children's theory of mind varied as a function of children's initial theory of mind level and degree of social isolation.

We evaluated the improvement of the goodness-of-fit of our model through each step using a nested model comparison approach based on the *deviance* (Hox, 2010). Models with a lower deviance fit better than models with a higher deviance. The difference between deviances for two nested models can be tested using a chi-square test, with degrees of freedom equal to the difference in the number of parameters estimated in the two models (Hox, 2010). Where there are non-significant differences between two models, the simpler model is preferred. Restricted Maximum Likelihood estimation (REML) estimates the fixed effects and the variance components separately and does not permit nested model comparisons (Field, 2009). REML is preferable when the number of clusters (i.e., classrooms) is fewer than 30 (McNeish & Stapleton, 2016). It provides unbiased fixed-effects point estimates for both level-1 and level-2 predictors (including cross-level interactions) with as few as 15 clusters, and unbiased estimates of level-2 variance components and standard errors of fixed-effects when performed with the Kenward-Roger correction (Bell, Morgan, Schoeneberger, Kromrey, & Ferron, 2014), even with as few as 4 clusters (Ferron, Bell, Hess, Rendina-Gobioff, & Hibbard, 2009). We tested our models using Maximum Likelihood to derive

the deviance statistics to compare models. We then used REML with a Kenward-Rogers correction to derive and report parameters.

To address the within-domain *specificity* of any contextual effects of peers' theory of mind on children's theory of mind, we evaluated a parallel model focused on children's verbal ability (instead of children's theory of mind). This model included children's verbal ability at Time 2 as the outcome variable and peers' theory of mind at Time 1 (i.e., the median and IQR) as predictors, along with children's verbal ability and theory of mind scores at Time 1 and the same control variables included in M2 described above.

Results

Preliminary analyses

Descriptive statistics and zero-order correlations are presented in Table 1 and Table 2, respectively for both child-level and classroom-level variables. The theory-of-mind latent factor showed partial scalar longitudinal invariance (Supplementary Material). To evaluate latent mean growth over time, we constrained the theory-of-mind latent mean to equality across time, by fixing both Time 1 and Time 2 means to zero. The model fit significantly decreased, $\Delta\text{CFI} = -.520$, $\Delta\text{RMSEA} = .046$, $\Delta\chi^2(1) = 283.153$, $p < .001$, indicating that the theory of mind latent increased significantly across time. Children's latent theory of mind score at Time 2 was, on average, 1.7 standard deviations higher than at Time 1.

Classroom-Related Contextual Effects on Individual's Theory of mind Development

Unstandardized parameter estimates and model fit statistics are presented in Table 3. Results of the Random Intercepts Empty model (M0) showed significant variation in random intercepts, $\text{var}(u_0) = 0.037$, $\Delta\chi^2(1) = 19.001$, $p < .001$, indicating that mean-level theory of mind scores at Time 2 varied across classrooms. Belonging to a particular class accounted for 11.1% of variation in children's theory of mind at Time 2. Before testing the effect of classroom-related contextual variables (i.e., median and IQR of peers' theory of mind), we estimated a Random Intercepts model (M1) to account for stability in children's theory of mind over time and for children's

characteristics (i.e., age, gender, verbal ability). The deviance statistic decreased moving from the Random Intercepts Empty model (M0) to this first Random Intercepts (M1) model, $\Delta\chi^2(8) = 152.108, p < .001$. Children's theory of mind at Time 2 was predicted by Time 1 theory of mind and by children's age, verbal ability and gender, with girls outperforming boys. Marginal R^2 (variance explained by only fixed effects –(Nakagawa & Schielzeth, 2013)) indicated that these child-level predictors accounted for 28.6% of variation in theory of mind at Time 2.

A second Random Intercept model (M2) was used to test whether classroom peers' theory of mind at Time 1 (i.e., median and IQR of peers' theory of mind) predicted variation in children's theory of mind at Time 2, over and above control variables included at the previous step. The deviance statistic decreased from M1 to M2, $\Delta\chi^2(2) = 7.958, p = .02$. There was a positive effect of diversity in peers' theory of mind at Time 1, $\beta = .14, p < .001$, but no effect of average levels of peers' theory of mind. Classroom theory of mind variety at Time 1 uniquely accounted for 1.1% of variation in children's theory of mind at Time 2.

We tested two models to examine if the strength of effect of peers' theory of mind varied depending on children's Time 1 theory of mind (M3a) or social isolation (M3b). The deviance statistic did not decrease moving from the more parsimonious Random Intercepts model (M2) to any of the two Interaction models, $\Delta\chi^2(3) < 2.420, ps > .49$. Results of each interaction model showed nonsignificant moderation effects ($ps > .33$). The effect of variation in classroom peers' theory of mind was comparable across varying levels of children's initial theory of mind scores at Time 1 and for children indexed as socially isolated and those who were not.

Specificity of Classroom-Related Contextual Effects

To examine the specificity of contextual effects of classroom peers' theory of mind on children's theory of mind at Time 2, we tested a parallel Random Intercept model in which children's verbal ability at Time 2 was regressed onto peers' theory of mind at Time 1 (i.e., median and IQR), verbal ability and children's theory of mind at Time 1. There was no significant contextual effect of peers' theory of mind on children's verbal ability.

Discussion

This study was designed to investigate, for the first time, the effect of peers' theory of mind on children's own theory of mind during middle childhood. Using multi-level longitudinal data, we examined whether the average level or variety of classroom peers' theory of mind predicted change in children's theory of mind over one year. We also tested whether classroom peer effects were moderated by children's own initial levels of theory of mind and social isolation and whether contextual effects of peers' theory of mind were specific to children's theory of mind or exerted more general effects on children's verbal ability. Our multi-level longitudinal analysis revealed three key findings. First, the variety (but not the average level) of classroom peers' theory of mind predicted children's theory of mind over time. This effect, although small, was significant even when a number of potential confounds at the individual level (i.e., age, gender, verbal ability, social isolation, special education needs, socio economic status) and initial levels of theory of mind were considered. Second, the effect of classroom peers' theory of mind was consistent regardless of children's initial levels of theory of mind or peer social isolation. Third, peers' theory of mind exhibited within-domain effects on children's theory of mind only and was not related to children's later verbal ability.

The current study makes an important contribution to research on theory of mind because, while social-cultural approaches to children's theory of mind development have a long tradition in developmental psychology (Hughes & Devine, 2015), existing studies have focused largely on family processes and theory of mind in young children. This is surprising for two reasons. First, children spend a large part of their day outside the family with classroom peers, meaning that a key source of social experience has been largely overlooked (McNeish & Stapleton, 2016). Second, twin studies show that the importance of nonshared environmental factors (such as classroom peers) on theory of mind increases with age (Ronald, Viding, Happé, & Plomin, 2006). Drawing on research on peer influences on children's academic skills (Yang et al., 2023), we tested two non-mutually exclusive hypotheses: a quantity and a variety hypothesis. Whereas the quantity

hypothesis focuses on the overall mean level of peers' theory of mind, the variety hypothesis highlights the importance of the variability in classroom peers' theory of mind. Supporting the variety hypothesis, belonging to a classroom with greater diversity in peers' theory of mind conferred benefits on children's own theory of mind over-and-above children's initial level of theory of mind, age, gender, SES, verbal ability, social isolation, presence of children with special education needs, and classroom size.

Our results are consistent with a small but growing body of research showing how peers contribute to the development of children's understanding of mind (Banerjee et al., 2011; Fink, Begeer, Peterson, Slaughter, & de Rosnay, 2015). Our findings extend this research by focusing on peers' theory of mind. The link between the variety in classroom peers' theory of mind and children's own theory of mind can be explained by existing theoretical accounts of cognitive development (Piaget, 1977; Vygotsky, 1978) and previous work on classroom influences on children's theory of mind (Smogorzewska et al., 2020). Together this work suggests that being exposed to diverse others is important for the ongoing refinement of theory of mind in middle childhood and early adolescence. Future work focused on the social interactions between skilled and less skilled mindreaders may shed light on the processes by which diversity in peers' theory of mind might benefit children's theory of mind.

We did not find a unique effect of the average level of classroom peers' theory of mind on children's theory of mind. This was somewhat unexpected given that longitudinal and training studies have shown that greater exposure to mental state conversations, for example, is linked with better theory of mind performance (Devine & Hughes, 2018). However, these findings echo those of existing research on the link between parents' and teachers' own theory of mind and children's theory of mind development. Specifically, Lecce and colleagues found that teachers' theory of mind ability was not significantly associated with pupils' own theory of mind (Lecce et al., 2022). Whether individual social partners have high or low levels of theory of mind ability might not

matter for children's own theory of mind as much as having social partners with diverse levels of theory of mind.

It is plausible that variety in the level of classroom peers' theory of mind increases children's exposure to different perspectives on the same reality, prompting mental state conversations and stimulating increased awareness of the subjective nature of mental states, and of the links between mental states and social behaviors. Different interpretations about the same classroom experiences may prompt children to reflect on and update their understanding of mental states, and consequently their ability to tune in to others' minds. Previous work suggests that differences in perspective can help children learn to negotiate, compromise, persuade, and take turns (Katz, Kramer, & Gottman, 1992), skills that may be underpinned by theory of mind. Beyond direct effects of peer social interactions, the influence of classroom peer diversity might be mediated via teachers, potentially affecting teaching strategies. Teachers of classes with greater diversity in theory-of-mind abilities might need to adjust their teaching strategies, tailoring them to meet diverse needs. For example, students with a well-developed theory of mind may comprehend abstract concepts and understand the perspectives of others more easily than their peers. Those with a less refined theory of mind might require more concrete explanations and support in understanding social dynamics. Recognizing these differences, teachers of classroom with greater variety in ToM might provide different interpretations of a given topic to cater to children's differing abilities. Consistent with this view, children's training (Bianco & Lecce, 2016) and observational studies (Lecce et al., 2022) demonstrate that teacher-prompted classroom conversations can improve children's theory of mind (Ily & Lai, 2011) and metacognition (Lecce, Demicheli, Zocchi, & Palladino, 2015). Direct observation of teachers' classroom interactions in settings with high and low diversity in theory of mind will illuminate the mechanisms involved.

The positive association between diversity in classroom peers' theory of mind and children's theory of mind was not moderated by children's own level of theory mind indicating that both initially skilled and unskilled mindreaders may benefit from being in a diverse classroom. Naturally

occurring individual differences in theory of mind during middle childhood and early adolescence (e.g., Devine & Lecce, 2021) benefit both skilled and less skilled members of a peer group but may do so through different mechanisms. For less skilled peers, the presence of more competent peers might enhance theory of mind through informal scaffolding opportunities because more skilled children may adopt the role of teachers (Lillard, 2005; Strauss & Ziv, 2012; Wellman & Lagattuta, 2004). For skilled children, interactions with less skilled peers might provide experience for refining their repertoire (i.e., their knowledge of what others know or do not know) (Ziv & Frye, 2004).

The positive effect of the variety of classroom peers' theory of mind on children's theory of mind was not moderated by children's social exclusion suggesting that children with fewer friends or relationships in the classroom can still benefit from classroom peers' theory of mind. One explanation is that isolated children, although having fewer opportunities to learn from direct interactions, can still benefit from observation (Akhtar, 2005; Akhtar, Jipson, & Callanan, 2001). This view fits with a training study by Gola (2012) who showed that preschoolers' false belief understanding can be enhanced through observing others' conversational exchanges. Peer effects on theory of mind may therefore be both direct and indirect.

The association between peers' variety in theory of mind and children's own theory of mind across time was domain-specific. The effect of peers' theory of mind did not extend to children's verbal ability, suggesting that the mechanisms linking peers' theory of mind and children's theory of mind may be domain-specific. Here it is important to note that the research on children's language development shows a significant and robust peer effect with the majority of these studies reporting that the average ability level of the peers in a child's classroom has positive effects on the child's language development after pretest scores and other control variables were controlled (Atkins-Burnett, Xue, & Aikens, 2017; Henry & Rickman, 2007; Justice et al., 2011; Mashburn et al., 2009). Future work examining within- and cross-domain effects of peers' theory of mind on other related cognitive (e.g., executive function) and academic abilities (e.g., reading) is needed.

Caveats and Conclusions

Some limitations deserve note. First, the relatively small number of classrooms included in the current study may have limited the power to detect cross-level interactions (Arend & Schafer, 2019). The small number of classrooms also prevented us from addressing the uniqueness of the effect of the classroom-level variety of peers' theory of mind on children's theory of mind. Although we controlled for several potential confounds at individual level, future research should address whether the effect of peers' theory of mind remains significant when accounting for variation in other classroom characteristics, such as verbal ability, socio-economic status¹ and ethnic composition. This issue is important as it is possible that multiple dimensions of diversity within classrooms, as reflected by variations in theory of mind, concurrently contribute to enhancing children's own theory of mind. A recent study, for example, has shown the importance of classroom ethnic diversity demonstrating its role in providing opportunities for cross-ethnicity friendships, which in turn promote children's own theory of mind abilities (controlling for verbal ability, executive function, peer social preference, and teacher-reported demographic characteristics) (Devine et al., 2024). The small number of classrooms also prevented us from addressing possible interactions between the variety and the amount hypothesis. Future research is needed to delve into the intricate interplay among diverse classroom factors and to investigate its role in children's theory of mind. Such endeavors will not only enrich our understanding of the mechanisms underpinning ToM development but also elucidate how classroom contexts can be optimized to cultivate children's socio-cognitive skills more effectively. Third, since we were not able to observe children's behavior in the classroom, the present study cannot shed light on the mechanisms underlying the effect of peer theory of mind variety. Notwithstanding these limitations, our results extend existing literature in three important ways. First, we showed that the classroom

¹ In response to a request from one reviewer we performed a follow up analysis adding SES and language at level 2. This analysis revealed that both the variety in SES, $\beta = .17, p = .050$, and the variety in verbal ability, $\beta = .34, p = .024$, were significant predictors of theory of mind development when considered separately from each other. We did not have sufficient power to enter all the predictors simultaneously.

environment shapes the development of children's theory of mind. This is important because peers (unlike family members) do not share genes raising the possibility of environmentally-mediated effects on theory of mind development. Second, the present study showed that social accounts of theory of mind can be extended to middle childhood and early adolescence. This is relevant because it helps researchers in building an age-appropriate model of the nature of social influence on further development and refinement in theory of mind. Third, our study highlights the importance of participation in diverse classrooms on children's theory of mind with possible implications for education policy, highlighting the advantage of mixing children with different abilities in the same classroom.

To our knowledge, the current study marks the first attempt to examine the effect of classroom peers' theory of mind on children's own theory of mind. Using a one-year longitudinal design and controlling for possible confounds, our study provides convincing evidence that diversity in peers' theory of mind can foster children's understanding of others' minds and sheds light on the processes by which theory of mind changes in middle childhood and early adolescence.

References

- Aikens, N., Tarullo, L., Hulseley, L., Ross, C., ... J. W.-A. for C., & 2010, undefined. (2010). A Year in Head Start: Children, Families and Programs. ACF-ORPRE Report. *ERIC*. Retrieved from <https://eric.ed.gov/?id=ED517213>
- Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, 8(2), 199–209. <https://doi.org/10.1111/j.1467-7687.2005.00406.x>
- Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, 72(2), 416–430. <https://doi.org/10.1111/1467-8624.00287>
- Andersen, A., Krølner, R., Currie, C., Dallago, L., Due, P., Richter, M., ... Holstein, B. E. (2008). High agreement on family affluence between children's and parents' reports: International study of 11-year-old children. *Journal of Epidemiology and Community Health*, 62(12), 1092–1094. <https://doi.org/10.1136/jech.2007.065169>

- Atkins-Burnett, S., Xue, Y., & Aikens, N. (2017). Peer Effects on Children's Expressive Vocabulary Development Using Conceptual Scoring in Linguistically Diverse Preschools. *Early Education and Development, 28*(7), 901–920.
<https://doi.org/10.1080/10409289.2017.1295585>
- Banerjee, R., Watling, D., & Caputi, M. (2011). Peer relations and the understanding of Faux Pas: Longitudinal evidence for bidirectional associations. *Child Development, 82*(6), 1887–1905.
<https://doi.org/10.1111/j.1467-8624.2011.01669.x>
- Bell, B. A., Morgan, G. B., Schoeneberger, J. A., Kromrey, J. D., & Ferron, J. M. (2014). How Low Can You Go? *https://Doi.Org/10.1027/1614-2241/A000062, 10*(1), 1–11.
<https://doi.org/10.1027/1614-2241/A000062>
- Bianco, F., & Lecce, S. (2016). Translating child development research into practice: Can teachers foster children's theory of mind in primary school? *British Journal of Educational Psychology, 86*(4). <https://doi.org/10.1111/bjep.12125>
- Brown, T. (2015). *Confirmatory factor analysis for applied research*.
<https://doi.org/10.1680/geot.8.B.012>
- Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics, 31*(1), 51–82. <https://doi.org/10.1086/666653>
- Carpendale, J., & Lewis, C. (2006). How children develop social understanding. Retrieved from Blackwell Publishing website: <https://psycnet.apa.org/record/2006-07153-000>
- Cassidy, K. W., Fineberg, D. S., Brown, K., & Perkins, A. (2005). Theory of mind may be contagious, but you don't catch it from your twin. *Child Development, 76*(1), 97–106.
<https://doi.org/10.1111/j.1467-8624.2005.00832.x>
- Cheung, E. O., & Gardner, W. L. (2015). The way I make you feel: Social exclusion enhances the ability to manage others' emotions. *Journal of Experimental Social Psychology, 60*, 59–75.
<https://doi.org/10.1016/j.jesp.2015.05.003>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing

measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.

https://doi.org/10.1207/S15328007SEM0902_5

Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age perspective. *Developmental Psychology*, 18(4), 557–570. <https://doi.org/10.1037/0012-1649.18.4.557>

Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: the development of the Health Behaviour in School-Aged Children (HBSC) family affluence scale. *Soc Sci Med*, 66(6), 1429–1436.

<https://doi.org/10.1016/j.socscimed.2007.11.024>

Deneault, J., Ricard, M., Decarie, T. G., Morin, P. L., Quintal, G., Boivin, M., ... Perusse, D. (2008). False belief and emotion understanding in monozygotic twins, dizygotic twins and non-twin children. *Cognition and Emotion*, 22(4), 697–708.

<https://doi.org/10.1080/02699930701486896>

Devine, R.T., & Apperly, I. A. (2022). Willing and able? Theory of mind, social motivation, and social competence in middle childhood and early adolescence. *Developmental Science*, 25(1).

<https://doi.org/10.1111/desc.13137>

Devine, R. T., Kovatchev, V., Grumley Traynor, I., Smith, P., & Lee, M. (2023). Machine learning and deep learning systems for automated measurement of “advanced” theory of mind: Reliability and validity in children and adolescents. *Psychological Assessment*, 35, 165–177.

<https://doi.org/10.1037/pas0001186>

Devine, R.T., & Lecce, S. (2021). Theory of mind in middle childhood and adolescence: Integrating multiple perspectives. In *Theory of Mind in Middle Childhood and Adolescence: Integrating Multiple Perspectives*. <https://doi.org/10.4324/9780429326899>

Devine, R.T., & Hughes, C. (2013). Silent Films and Strange Stories: Theory of Mind, gender, and social experiences in middle childhood. *Child Development*, 84(3), 989–1003.

<https://doi.org/10.1111/cdev.12017>

- Devine, R.T., & Hughes, C. (2016). Measuring theory of mind across middle childhood: Reliability and validity of the Silent Films and Strange Stories tasks. *Journal of Experimental Child Psychology, 149*, 23–40. <https://doi.org/10.1016/J.JECP.2015.07.011>
- Devine, R.T., & Hughes, C. (2017). Let's talk: Parents' mental talk (not mind-mindedness or mindreading capacity) predicts children's false belief understanding. *Child Development*. <https://doi.org/10.1111/cdev.12990>
- Devine, R.T., & Hughes, C. (2018). Family Correlates of False Belief Understanding in Early Childhood: A Meta-Analysis. *Child Development, 89*(3), 971–987. <https://doi.org/10.1111/cdev.12682>
- Devine, R.T., Traynor, I.G., Ronchi, L., & Lecce, S. (2024). Children in ethnically diverse classrooms and those with cross-ethnic friendships excel at understanding others' minds. To appear in *Child Development*. <https://doi.org/10.1111/cdev.14085>
- Ferron, J. M., Bell, B. A., Hess, M. R., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: The utility of multilevel modeling approaches. *Behavior Research Methods, 41*(2), 372–384. <https://doi.org/10.3758/BRM.41.2.372>
- Field, A. (2009). *Discovering Statistics Using SPSS*. In Sage Publication.
- Fink, E., Begeer, S., Peterson, C. C., Slaughter, V., & de Rosnay, M. (2015). Friendlessness and theory of mind: A prospective longitudinal study. *British Journal of Developmental Psychology, 33*(1), 1–17. <https://doi.org/10.1111/bjdp.12060>
- Gola, A. A. H. (2012). Mental verb input for promoting children's theory of mind: A training study. *Cognitive Development, 27*(1), 64–76. <https://doi.org/10.1016/j.cogdev.2011.10.003>
- Hanushek, E., Kain, J., ... J. M.-J. of applied, & 2003, undefined. (2003). Does peer ability affect student achievement? *Wiley Online Library* EA Hanushek, JF Kain, JM Markman, SG Rivkin *Journal of Applied Econometrics, 2003*•Wiley Online Library, 18(5), 527–544. <https://doi.org/10.1002/jae.741>

- Happé, F. G. E., & Happe, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*(2), 129–154.
<https://doi.org/10.1007/BF02172093>
- Henry, G. T., & Rickman, D. K. (2007). Do peers influence children's skill development in preschool? *Economics of Education Review*, *26*(1), 100–112.
<https://doi.org/10.1016/J.ECONEDUREV.2005.09.006>
- Hofmann, S. G., Doan, S. N., Sprung, M., Wilson, A., Ebesutani, C., Andrews, L. A., ... 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>
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. Routledge.
- Hughes, C., & Devine, R. T. (2015). Individual Differences in Theory of Mind From Preschool to Adolescence: Achievements and Directions. *Child Development Perspectives*, *9*(3), 149–153.
<https://doi.org/10.1111/cdep.12124>
- Hughes, C., & Devine, R. T. (2015). A Social Perspective on Theory of Mind. In *Handbook of Child Psychology and Developmental Science* (pp. 1–46).
<https://doi.org/10.1002/9781118963418.childpsy314>
- Hughes, C., & Devine, R. T. (2019). For Better or for Worse? Positive and Negative Parental Influences on Young Children's Executive Function. *Child Development*, *90*(2), 593–609.
<https://doi.org/10.1111/cdev.12915>
- Ily, E., & Lai, R. (2011). *Metacognition: A Literature Review Research Report*. Retrieved from <http://www.pearsonassessments.com/research>.
- Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, *32*(1), 70–78.
<https://doi.org/10.1037/0012-1649.32.1.70>
- Jenkins, J. M., Turrell, S. L., Kogushi, Y., Louis, S., & Ross, H. S. (2003). A Longitudinal

- Investigation of the Dynamics of Mental State Talk in Families. *Child Development*, 74(3), 905–920. <https://doi.org/10.1111/1467-8624.00575>
- Justice, L. M., Petscher, Y., Schatschneider, C., & Mashburn, A. (2011). Peer Effects in Preschool Classrooms: Is Children's Language Growth Associated With Their Classmates' Skills? *Child Development*, 82(6), 1768–1777. <https://doi.org/10.1111/j.1467-8624.2011.01665.x>
- Katz, L. F., Kramer, L., & Gottman, J. M. (1992). Conflict and emotions in marital, sibling, and peer relationships. In C. U. Hartup & W. W. Shantz (Eds.), *Conflict in child and adolescent development* (pp. 122–149). Cambridge University Press.
- Kennedy, K., Lagattuta, K. H., & Sayfan, L. (2015). Sibling composition, executive function, and children's thinking about mental diversity. *Journal of Experimental Child Psychology*, 132, 121–139. <https://doi.org/10.1016/J.JECP.2014.11.007>
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151–160. <https://doi.org/10.1007/BF02288391>
- Lecce, S. (2021). Theory of mind goes to school. In *Theory of Mind in Middle Childhood and Adolescence: Integrating Multiple Perspectives*.
- 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. <https://doi.org/10.1016/j.jecp.2014.03.002>
- Lecce, S., Bianco, F., & Hughes, C. (2021). Reading minds and reading texts: Evidence for independent and specific associations. *Cognitive Development*, 57. <https://doi.org/10.1016/j.cogdev.2021.101010>
- Lecce, S., Demicheli, P., Zocchi, S., & Palladino, P. (2015). The origins of children's metamemory: The role of theory of mind. *Journal of Experimental Child Psychology*, 131. <https://doi.org/10.1016/j.jecp.2014.11.005>
- Lecce, S., Ronchi, L., & Devine, R. T. (2022). Mind what teacher says: Teachers' propensity for mental-state language and children's theory of mind in middle childhood. *Social Development*,

31(2), 303–318. <https://doi.org/10.1111/sode.12552>

Lillard, A. (2005). *Montessori: The Science Behind the Genius*. Oxford University Press.

Mashburn, A. J., Justice, L. M., Downer, J. T., & Pianta, R. C. (2009). Peer effects on children's language achievement during pre-kindergarten. *Child Development, 80*(3), 686–702. <https://doi.org/10.1111/j.1467-8624.2009.01291.x>

McNeish, D. M., & Stapleton, L. M. (2016). The Effect of Small Sample Size on Two-Level Model Estimates: A Review and Illustration. *Educational Psychology Review, 28*(2), 295–314. <https://doi.org/10.1007/s10648-014-9287-x>

Merlo, J., Wagner, P., Austin, P. C., Subramanian, S. V., & Leckie, G. (2018). General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: A conceptual tutorial. *SSM - Population Health, 5*, 33–37. <https://doi.org/10.1016/j.ssmph.2018.05.006>

Muthèn, L. K., & Muthén, B. O. (2017). *Mplus: Statistical Analysis With Latent Variables. User's Guide* (8th ed.; Muthèn & Muthén., Ed.).

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>

Osterhaus, C., & Koerber, S. (2023). The complex associations between scientific reasoning and advanced theory of mind. *Child Development, 94*(1), e18–e42. <https://doi.org/10.1111/cdev.13860>

Parker, J. G., & Asher, S. R. (1993). Friendship and friendship quality in middle childhood: Links with peer group acceptance and feelings of loneliness and social dissatisfaction. *Developmental Psychology, 29*(4), 611–621. <https://doi.org/10.1037/0012-1649.29.4.611>

Peterson, C. C. (2000). Kindred spirits: Influences of siblings' perspectives on theory of mind. *Cognitive Development, 15*(4), 435–455. [https://doi.org/10.1016/s0885-2014\(01\)00040-5](https://doi.org/10.1016/s0885-2014(01)00040-5)

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures.* (Trans A.

Rosin). Retrieved from <https://psycnet.apa.org/record/1979-20791-000>

- Ronald, A., Viding, E., Happé, F., & Plomin, R. (2006). Individual differences in theory of mind ability in middle childhood and links with verbal ability and autistic traits: a twin study. *Social Neuroscience, 1*(3–4), 412–425. <https://doi.org/10.1080/17470910601068088>
- Rubini, V., & Rossi, M. A. (1982). Analysis of the psychometric characteristics of the Primary Mental Abilities Tests (Level 4-sup-6, 1962 revision) applied to a group of Italian subjects. *Bollettino Di Psicologia Applicata, 161*, 87–105.
- Rust, J. (2008). *Raven's Standard Progressive Matrices and Mill Hill Vocabulary Scale*. Pearson Education.
- Schnohr, C. W., Kreiner, S., Due, E. P., Currie, C., Boyce, W., & Diderichsen, F. (2008). Differential item functioning of a family affluence scale: Validation study on data from HBSC 2001/02. *Social Indicators Research, 89*(1), 79–95. <https://doi.org/10.1007/s11205-007-9221-4>
- Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-Analysis of Theory of Mind and Peer Popularity in the Preschool and Early School Years. *Child Development, 86*(4), 1159–1174. <https://doi.org/10.1111/cdev.12372>
- Smogorzewska, J., Szumski, G., & Grygiel, P. (2020). Theory of mind goes to school: Does educational environment influence the development of theory of mind in middle childhood? *PLoS ONE, 15*(8 August). <https://doi.org/10.1371/journal.pone.0237524>
- Strauss, S., & Ziv, M. (2012). Teaching is a natural cognitive ability for humans. *Mind, Brain, and Education, 6*(4), 186–196. <https://doi.org/10.1111/j.1751-228X.2012.01156.x>
- Thurstone, T. G., & Thurstone, L. L. (1962). *Primary mental abilities tests*. Science Research Associates.
- Torsheim, T., Cavallo, F., Levin, K. A., Schnohr, C., Mazur, J., Niclasen, B., & Currie, C. (2016). Psychometric Validation of the Revised Family Affluence Scale: a Latent Variable Approach. *Child Indicators Research, 9*(3), 771–784. <https://doi.org/10.1007/s12187-015-9339-x>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*.

Cambridge, MA: Harvard University Press.

Wang, Y., & Su, Y. (2009). False belief understanding: Children catch it from classmates of different ages. *International Journal of Behavioral Development*, 33(4), 331–336.

<https://doi.org/10.1177/0165025409104525>

Wellman, H. M., & Lagattuta, K. H. (2004). Theory of mind for learning and teaching: The nature and role of explanation. *Cognitive Development*, 19(4), 479–497.

<https://doi.org/10.1016/J.COGDEV.2004.09.003>

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development*, 80, 1097–1117.

<https://doi.org/10.1111/j.1467-8624.2009.01319.x>

Yang, Q., Ansari, A., Purtell, K. M., Pianta, R. C., Whittaker, J. V., & Vitiello, V. E. (2023). Classroom skill compositions and preschoolers' early academic and executive function outcomes. *Early Childhood Research Quarterly*, 64, 290–301.

<https://doi.org/10.1016/j.ecresq.2023.04.003>

Ziv, M., & Frye, D. (2004). Children's understanding of teaching: The role of knowledge and belief. *Cognitive Development*, 19(4), 457–477.

<https://doi.org/10.1016/J.COGDEV.2004.09.002>

Table 1 *Descriptive Statistics and Zero-Order Correlations for Children-Related Variables*

	M (SD)	Min - Max	2	3	4	5
1. Age	9.43 (.68)	7.95 – 12.20	.03	.16**	.29***	.25
2. SES	6.56 (1.60)	1 – 9	-	.18***	.02	.04
3. VA	-0.03 (1.03)	-5.05 – 2.86		-	.29***	.32***
4. T1 ToM	1.82 (0.65)	0.25 – 3.50			-	.46***
5. T2 ToM	2.51 (0.58)	0.50 – 3.83				-

Note. Age = Children's age; SES = Children's socioeconomic status; VA = Children's verbal ability; ToM = Children's theory of mind; ** $p \leq .01$; *** $p \leq .001$.

Table 2 *Descriptive Statistics and Zero-Order Correlations for Classroom-Related Variables*

	M (SD)	Min - Max	2	3
1. Class Size	22.70 (3.42)	15 – 28	-.07	-.48*
2. Class ToM-Median	1.80 (0.28)	1.29 – 2.42	-	.21
3. ClassToM-IQR	0.87 (0.19)	0.58 – 1.31		-

Note. Class ToM-Median = Median level of classroom's theory of mind; Class ToM-RIQ =

Interquartile range of classroom's theory of mind* $p \leq .05$;

Table 3 *Intercept-only and nested models with explanatory variables*

Model	M1: Intercept-Only	M2: Within-Classroom	M3: Between-Classroom
Fixed Part	Coefficient (s.e.)	Coefficient (s.e.)	Coefficient (s.e.)
Intercept	2.51 (.04)	2.41 (0.5)	2.40 (0.5)
<i>Child-level variables</i>			
Age		.07 (.03) *	.06 (.06)
SES		-.02 (.03)	-.03 (.05)
VA		.09 (.03) **	.15 (.05) **
T1 TOM		.20 (.03) ***	.34 (.05) ***
Gender (M – F)		-.14 (.05) **	-.26 (.09) **
SEND (Yes – No)		-.16 (.10)	-.32 (.17) +
Social Isolation (Yes – No)		-.13 (.09)	-.24(.16)
<i>Classroom-level variables</i>			
Class-Size		.01 (.03)	.08 (.05)
T1 class-TOM level			.07 (.06)
T1 class-TOM variety			.14 (.05) **
Random Part			
σ^2_e	0.297	0.222	0.714
σ^2_{u0}	0.037	0.006	0.000
Model Statistics			
Deviance (-2logLL)	614.314	462.206	454.248
ICC	0.111	.026	-

Note. Age = Children's age; SES = Children's socioeconomic status; VA = Children's verbal ability; T1 ToM = Time 1 ToM; SEND = Special education needs/disabilities. Reported model

parameters are standardized estimates, except for the intercept (unstandardized score). ⁺ $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$.