



To the Roots of Theory of Mind Deficits in Autism Spectrum Disorder: A Narrative Review

Katarína Polónyiová¹ · Joanna Kruyt^{2,3} · Daniela Ostatníková¹

Received: 9 February 2024 / Accepted: 21 March 2024
© The Author(s) 2024, corrected publication 2024

Theory of Mind

Theory of mind (ToM) is a term first used by Premack and Woodruff, who defined it as the ability of an individual to attribute different mental states to themselves as well as to others (Premack & Woodruff, 1978). Apperly and Butterfill (2009) added that ToM includes the awareness that the mental states (beliefs, desires, intentions, emotions, and thoughts) of other people can be different from our own and they can be used to explain or predict the behavior of others. ToM deficits are traditionally associated with individuals with autism spectrum disorder (ASD) and can represent in different degrees from complete “mindblindness” to mastering even higher tests, but still struggling in practical social situations (Baron-Cohen, 1995).

It is important to note that ToM is not an isolated cognitive skill, and it is likely interconnected with other cognitive functions such as intelligence, executive functions or language. The mutual connections between these factors have been extensively researched, yet the precise causal mechanisms remain elusive.

The main objectives of this paper are therefore to summarize previously published findings about ToM of individuals

with ASD in context of their intelligence, executive functions, language development, and the double empathy problem, to elaborate on competing perspectives and identify possible research gaps. Moreover, we aim to propose our recommendations for future research and clinical practice.

Theory of Mind and Autism Spectrum Disorder

Baron-Cohen (1995) considered ToM deficits to be the underlying cause behind the core symptoms of ASD, a neurodevelopmental disorder characterized by persistent deficits in the ability to initiate and to sustain reciprocal social interaction and social communication, and by a range of restricted, repetitive, and inflexible patterns of behavior and interests (World Health Organisation, 2019). It has been shown that the ToM deficits severity can reliably distinguish the severity level of the core symptoms of ASD (Hoogenhout & Malcolm-Smith, 2017), but there is no consensus yet on whether ToM is inherently deficient in individuals with ASD, or its development is only significantly delayed.

The acquisition of ToM in neurotypical children typically follows a certain order. Understanding of intentions develops first, then the ability to understand desires, and finally the ability to understand beliefs. In contrast, studies that investigate the developmental trajectory of ToM in children with ASD do not paint as clear a picture. Broekhof et al. (2015) found that children with ASD and neurotypical children have the same ability to understand intentions, but children with ASD have greater difficulty understanding desires and beliefs. Other authors observed that children with ASD have deficits in understanding intentions and beliefs, but do not differ in understanding emotions (Sivaratnam et al., 2012), or deficits in understanding emotions and beliefs, but no impairment in understanding intentions (Mazza et al., 2017), or even deficits in understanding beliefs and emotions, with intact understanding of intentions (Pino et al., 2017).

✉ Katarína Polónyiová
katarina.polonyiova@gmail.com;
katarina.polonyiova@fmed.uniba.sk

Joanna Kruyt
joannakruyt@gmail.com

Daniela Ostatníková
daniela.ostatnikova@fmed.uniba.sk

¹ Institute of Physiology, Academic Research Centre for Autism, Faculty of Medicine, Comenius University in Bratislava, Sasinkova 2, 813 72 Bratislava, Slovakia

² Institute of Informatics, Slovak Academy of Sciences in Bratislava, Bratislava, Slovakia

³ Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava, Bratislava, Slovakia

Different researchers hypothesize that the order in which these aspects of ToM develop is identical in children with ASD, but usually significantly delayed. Four- to five-year-old neurotypical children show the first signs of ToM, which become more sophisticated over time, and by around the age of 6 or 7 years, they can reliably pass the highest level of ToM evaluation—the false belief tests (Happé & Frith, 2014). It appears that a minimum mental age of 11 years is required for a child with ASD to have at least an 80% chance of passing such tests (Happé, 1995). According to some studies, ToM in individuals with ASD can continue improving even during middle childhood and adolescence (Pino et al., 2017; Steele et al., 2003).

Theory of Mind and Intelligence

Most psychometric models consider intelligence to be a latent variable, i.e., a general ability that plays a causal role in a wide range of cognitive tasks, including ToM. Psychometric models have shown that individual differences in ToM can be largely attributed to general intelligence (g-factor) (Coyle et al., 2018), meaning that some IQ threshold is likely essential for the development of ToM, and higher IQ is generally associated with better ToM abilities. One of the most commonly used tests of ToM, the Reading the Mind in the Eyes Test, created by Baron-Cohen, has been used in many studies along with IQ tests, and the majority of these studies showed a positive correlation between ToM and IQ, with no differences between verbal and nonverbal IQ tests (Baker et al., 2014). Alternative perspectives argue that the relationship between IQ and ToM is mediated by other variables, such as executive functions or language abilities (Navarro et al., 2021). Rajkumar et al. (2008) found no relationship between ToM and IQ after controlling for confounding variables such as general language ability and working memory.

Theory of Mind and Executive Functions

Executive functions (EFs) are a set of higher-order cognitive processes which regulate a variety of abilities necessary for goal-directed behavior. There is an ongoing discussion about which processes can be included in EFs, but the most common division includes three basic EFs: working memory, inhibition, and shifting (Lehto et al., 2003).

The role of EFs in ToM deficits is especially complicated to define, since EFs are likely interconnected with all other cognitive processes. Above all, EFs are a similar and closely correlated concept to IQ and they partially overlap in some aspects (García-Molina et al., 2010). Executive dysfunction is very common in individuals with ASD, and even

children and adolescents with average intellect often struggle with cognitive flexibility, impulse inhibition, or overall problem-solving (Merchán-Naranjo et al., 2016). Kouklari et al. (2018) found that children with ASD and neurotypical children only differed in inhibition and cognitive flexibility, but all EFs correlated with ToM and working memory best predicted ToM.

Moses (2001) proposed that the relationship between ToM and EFs can be explained by two main theories: according to the expression account, ToM tasks are so demanding and require such a high level of EFs that they may not allow children to fully express their true ToM abilities. Therefore, specific failures of some children on ToM tasks may reflect EF deficits rather than ToM deficits, and these children may be capable of ToM, but they are unable to demonstrate it in a given situation because their EF deficits prevent them from doing so. Some studies suggest a predictive relationship between EFs and ToM in both neurotypical children and children with ASD, regardless of age, and verbal or nonverbal IQ (Pellicano, 2010).

According to the emergence account, EFs are, on the contrary, a necessary condition for the development of ToM (Moses, 2001). Children need to be able to mentally work with different perspectives independently of their own perspective, and they must have developed and applied all EFs in order to do this. Adequately developed EFs could simply make it easier for children to complete ToM tasks, since the performance of the task itself requires a certain amount of EFs (Hughes & Ensor, 2007). The observed correlation between ToM and EF deficits in ASD may also be due to the fact that both of these abilities depend on the same or closely connected brain structures, and therefore damage to one area can cause problems in both EFs and ToM (Perner & Lang, 2000).

Theory of Mind and Language

Several studies suggest that there may be a relationship between children's performance on ToM tests and their performance on different language development assessments (Cutting & Dunn, 1999; Watson et al., 2001). The cause, as well as the direction of this strong association, are point of discussion.

Some researchers attribute the strong association to the observation that most ToM tests incorporate verbal elements, and these tests may thus in some way conflate an individual's level of language ability with their ToM development (Bloom & German, 2000). A similar explanation for the apparent link between ToM and language development also posits that the two are not related per se, but rather that they both rely on shared underlying cognitive resources that change and develop over time, such as EFs (Carlson & Moses, 2001). Indeed, in a

study by Shahaiean et al. (2023), language was found to influence the development of both ToM and EFs. Alternatively, it has been hypothesized that language development and ToM seem linked because children who hear adults speak about others' mental states develop both their ToM and their language abilities simultaneously. Hearing adults talk about other people's feelings may predict a child's ToM test score (Ruffman et al., 2002).

Other theories argue for a stronger and more explicit causal link between ToM and language development, postulating that a certain level of linguistic ability is directly required for the development of ToM (De Villiers & De Villiers, 2014) or that language has a mediating effect on the relationship of age and both cognitive and affective ToM (Bigelow et al., 2021). Moreover, the intervention study by Hale and Tager-Flusberg (2003) proved that the training focused on the acquisition of the syntactic and semantic properties significantly increased ToM of preschoolers, comparably to an intervention aimed directly at improvement of ToM.

Theory of Mind and the Double Empathy Problem

Finally, rather than attributing social and communicative difficulties experienced by individuals with ASD to inherent ToM deficits, it is proposed that these challenges may arise from "neurotype mismatches" occurring during interactions between individuals with ASD and neurotypical ones. Individuals with ASD may not inherently lack a ToM but may lack a theory of neurotypical mind, specifically. Importantly, this works both ways: neurotypical individuals also appear to lack a theory of autistic mind (Heasman & Gillespie, 2019; Sheppard et al., 2016). Termed the "double empathy problem" (Milton, 2012), this perspective is often endorsed by individuals with ASD. Recent research appears to support the double empathy problem. For example, Morrison et al. (2020) found that individuals with ASD seem to prefer having unstructured conversations with others who shared their diagnosis, rather than with neurotypical individuals. Moreover, higher levels of rapport (Crompton et al., 2020) and flow (Williams, 2021) have been reported in conversations between neurotype-matched dyads (i.e., ASD-ASD and neurotypical-neurotypical) than during interactions with a neurotype mismatch.

Conclusions and Future Directions for Research and Practice

Despite extensive research, the roots of ToM deficits observed in individuals with ASD remain elusive. The deficits could mainly be the result of the primary diagnosis of ASD or arise due to misunderstanding with neurotypical

peers. ToM is also closely linked to other cognitive processes (IQ, EFs, or language) which are not only conceptually interconnected, but also known for maturing during the same periods of children's development. In this review, we aimed to explore these existing debates and to summarize the fundamental evidence for all of the aforementioned theories.

It is important to note that the amount of research proving the relationship between EFs and ToM is significantly higher than for other variables, which would provide a hopeful perspective, as EFs are known to be relatively trainable. However, further research is necessary to prove this claim. Proving the generalizability of the conclusions of previous studies is challenging due to several factors: (a) separating between the roles of interrelated cognitive functions is difficult in the first place, (b) the majority of the previous studies focus only on the combinations of some of these factors thus (c) they cannot fully explore the relations between these mutually confounding variables.

Therefore, our main recommendation and direction for research in the future include assessing a wider range cognitive processes in a more prominent research sample, to help clarify their mutual relations. This should also lead to implementing more advanced methodological and statistical methods, like regression models, mediation, and modulation analyses which provide more reliable data on the directionalities of these relationships.

Finally, further involvement of autistic individuals is crucial for future research, to be able to consider the double empathy problem as well. This alternative theory surely does not mean the end of ToM research. On the contrary, we believe that the paradigm shift could lead to the creation of new methods and to the further development of ToM understanding, improved by incorporating the perspectives of the very individuals with ASD.

Funding This research was funded by the Slovak Research and Development Agency, grant numbers APVV-20-0070 and APVV 20-0139.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 116(4), 953–970. <https://doi.org/10.1037/a0016923>

- Baker, C. A., Peterson, E., Pulos, S., & Kirkland, R. A. (2014). Eyes and IQ: A meta-analysis of the relationship between intelligence and "Reading the Mind in the Eyes." *Intelligence*, *44*, 78–92. <https://doi.org/10.1016/j.intell.2014.03.001>
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. The MIT Press.
- Bigelow, F. J., Clark, G. M., Lum, J. A. G., & Enticott, P. G. (2021). The mediating effect of language on the development of cognitive and affective theory of mind. *Journal of Experimental Child Psychology*, *209*, 105158. <https://doi.org/10.1016/j.jecp.2021.105158>
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), B25–B31. [https://doi.org/10.1016/S0010-0277\(00\)00096-2](https://doi.org/10.1016/S0010-0277(00)00096-2)
- Broekhof, E., Ketelaar, L., Stockmann, L., van Zijp, A., Bos, M. G. N., & Rieffe, C. (2015). The understanding of intentions, desires and beliefs in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(7), 2035–2045. <https://doi.org/10.1007/s10803-015-2363-3>
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053. <https://doi.org/10.1111/1467-8624.00333>
- Coyle, T. R., Elpers, K. E., Gonzalez, M. C., Freeman, J., & Baggio, J. A. (2018). General intelligence (g), ACT scores, and theory of mind: (ACT)g predicts limited variance among theory of mind tests. *Intelligence*, *71*, 85–91. <https://doi.org/10.1016/j.intell.2018.10.006>
- Crompton, C. J., Ropar, D., Evans-Williams, C. V., Flynn, E. G., & Fletcher-Watson, S. (2020). Autistic peer-to-peer information transfer is highly effective. *Autism*, *24*(7), 1704–1712. <https://doi.org/10.1177/1362361320919286>
- Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, *70*(4), 853–865. <https://doi.org/10.1111/1467-8624.00061>
- De Villiers, J. G., & De Villiers, P. A. (2014). The role of language in theory of mind development. *Topics in Language Disorders*, *34*(4), 313–328. <https://doi.org/10.1097/TLD.0000000000000037>
- García-Molina, A., Tirapu-Ustárrroz, J., Luna-Lario, P., Ibáñez, J., & Duque, P. (2010). Are intelligence and executive functions the same thing? *Revista De Neurologia*, *50*(12), 738–746.
- Gioia, G. A., Isquith, P. K., Kenworthy, L., & Barton, R. M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*, *8*(2), 121–137. <https://doi.org/10.1076/chin.8.2.121.8727>
- Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science*, *6*(3), 346–359. <https://doi.org/10.1111/1467-7687.00289>
- Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843–855.
- Happé, F., & Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry*, *55*(6), 553–577. <https://doi.org/10.1111/jcpp.12162>
- Heasman, B., & Gillespie, A. (2019). Neurodivergent intersubjectivity: Distinctive features of how autistic people create shared understanding. *Autism*, *23*(4), 910–921. <https://doi.org/10.1177/1362361318785172>
- Hoogenhout, M., & Malcolm-Smith, S. (2017). Theory of mind predicts severity level in autism. *Autism*, *21*(2), 242–252. <https://doi.org/10.1177/1362361316636758>
- Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*(6), 1447–1459. <https://doi.org/10.1037/0012-1649.43.6.1447>
- Im-Bolter, N., Agostino, A., & Owens-Jaffray, K. (2016). Theory of mind in middle childhood and early adolescence: Different from before? *Journal of Experimental Child Psychology*, *149*, 98–115. <https://doi.org/10.1016/j.jecp.2015.12.006>
- Jones, C. R. G., Simonoff, E., Baird, G., Pickles, A., Marsden, A. J. S., Tregay, J., Happé, F., & Charman, T. (2018). The association between theory of mind, executive function, and the symptoms of autism spectrum disorder: Cognitive and behavioural associations in ASD. *Autism Research*, *11*(1), 95–109. <https://doi.org/10.1002/aur.1873>
- Kouklari, E.-C., Tsermentseli, S., & Auyeung, B. (2018). Executive function predicts theory of mind but not social verbal communication in school-aged children with autism spectrum disorder. *Research in Developmental Disabilities*, *76*, 12–24. <https://doi.org/10.1016/j.ridd.2018.02.015>
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80. <https://doi.org/10.1348/026151003321164627>
- Mazza, M., Mariano, M., Peretti, S., Masedu, F., Pino, M. C., & Valenti, M. (2017). The role of theory of mind on social information processing in children with autism spectrum disorders: A mediation analysis. *Journal of Autism and Developmental Disorders*, *47*(5), 1369–1379. <https://doi.org/10.1007/s10803-017-3069-5>
- Merchán-Naranjo, J., Boada, L., del Rey-Mejías, Á., Mayoral, M., Llorente, C., Arango, C., & Parellada, M. (2016). Executive function is affected in autism spectrum disorder, but does not correlate with intelligence. *Revista De Psiquiatría y Salud Mental (english Edition)*, *9*(1), 39–50. <https://doi.org/10.1016/j.rpsmen.2016.01.001>
- Milton, D. E. M. (2012). On the ontological status of autism: The 'double empathy problem.' *Disability & Society*, *27*(6), 883–887. <https://doi.org/10.1080/09687599.2012.710008>
- Morrison, K. E., DeBrabander, K. M., Jones, D. R., Faso, D. J., Ackerman, R. A., & Sasson, N. J. (2020). Outcomes of real-world social interaction for autistic adults paired with autistic compared to typically developing partners. *Autism*, *24*(5), 1067–1080. <https://doi.org/10.1177/1362361319892701>
- Moses, L. J. (2001). Executive accounts of theory-of-mind development. *Child Development*, *72*(3), 688–690. <https://doi.org/10.1111/1467-8624.00306>
- 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>
- Pellicano, E. (2010). Individual differences in executive function and central coherence predict developmental changes in theory of mind in autism. *Developmental Psychology*, *46*(2), 530–544. <https://doi.org/10.1037/a0018287>
- Perner, J., & Lang, B. (2000). *Theory of mind and executive function: Is there a developmental relationship?* Perspectives from developmental cognitive neuroscience. Oxford University Press.
- Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development*, *76*(2), 502–517. <https://doi.org/10.1111/j.1467-8624.2005.00859.x>
- Pino, M. C., Mazza, M., Mariano, M., Peretti, S., Dimitriou, D., Masedu, F., Valenti, M., & Franco, F. (2017). Simple mindreading abilities predict complex theory of mind: Developmental delay in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *47*(9), 2743–2756. <https://doi.org/10.1007/s10803-017-3194-1>
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. <https://doi.org/10.1017/S0140525X00076512>

- Rajkumar, A. P., Yovan, S., Raveendran, A. L., & Russell, P. S. S. (2008). Can only intelligent children do mind reading: The relationship between intelligence and theory of mind in 8 to 11 years old. *Behavioral and Brain Functions*, 4(1), 51. <https://doi.org/10.1186/1744-9081-4-51>
- Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, 73(3), 734–751. <https://doi.org/10.1111/1467-8624.00435>
- Shahaeian, A., Haynes, M., & Frick, P. J. (2023). The role of language in the association between theory of mind and executive functioning in early childhood: New longitudinal evidence. *Early Childhood Research Quarterly*, 62, 251–258. <https://doi.org/10.1016/j.ecresq.2022.09.003>
- Sheppard, E., Pillai, D., Wong, G.T.-L., Ropar, D., & Mitchell, P. (2016). How Easy is it to read the minds of people with autism spectrum disorder? *Journal of Autism and Developmental Disorders*, 46(4), 1247–1254. <https://doi.org/10.1007/s10803-015-2662-8>
- Sivaratnam, C. S., Cornish, K., Gray, K. M., Howlin, P., & Rinehart, N. J. (2012). Brief report: Assessment of the social-emotional profile in children with autism spectrum disorders using a novel comic strip task. *Journal of Autism and Developmental Disorders*, 42(11), 2505–2512. <https://doi.org/10.1007/s10803-012-1498-8>
- Steele, S., Joseph, R. M., & Tager-Flusberg, H. (2003). Brief report: Developmental change in theory of mind abilities in children with autism. *Journal of Autism and Developmental Disorders*, 33(4), 461–467. <https://doi.org/10.1023/A:1025075115100>
- Watson, A. C., Painter, K. M., & Bornstein, M. H. (2001). Longitudinal Relations between 2-year-olds' language and 4-year-olds' theory of mind. *Journal of Cognition and Development*, 2(4), 449–457. https://doi.org/10.1207/S15327647JCD0204_5
- Williams, G. L. (2021). Theory of autistic mind: A renewed relevance theoretic perspective on so-called autistic pragmatic 'impairment.' *Journal of Pragmatics*, 180, 121–130. <https://doi.org/10.1016/j.pragma.2021.04.032>
- World Health Organisation. (2019). *ICD-11 International Classification of Diseases 11th Revision*. <https://icd.who.int/en/>. Accessed 2 Feb 2024

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.