Psychology of Language and Communication 2019, Vol. 23, No. 1

DOI: 10.2478/plc-2019-0006

MACIEJ HAMAN
University of Warsaw, Warsaw, Poland

DEVELOPING THEORY OF MIND TWENTY-FIVE YEARS AFTER THE PUBLICATION OF "Z BADAŃ NAD KOMPETENCJĄ KOMUNIKACYJNĄ DZIECKA" (EDITED BY B. BOKUS AND M.HAMAN)

Twenty-five years ago, a book "Z badań nad kompetencją komunikacyjną dziecka", edited by Barbara Bokus and Maciej Haman, was issued containing, among else, the first Polish review of the studies on the development of Theory of Mind. During these 25 years, the area developed extensively and a new "state-of-the-arts" paper is necessary. The current paper does not pretend to the role of a complete review, instead it focusses on two live issues in the Theory of Mind (ToM) research progress: early (before the age of four years) competences in false-belief understanding, which leads to the question of continuity versus discontinuity (e.g., "Two-system theory") between early and later ToM abilities, and neuroimaging studies of Theory-of-Mind, which may also contribute to the continuity debate.

Keywords: theory of mind, implicit and explicit false belief task, cognitive development, neuroimaging

Twenty-five years ago, a book "Z badań nad kompetencją komunikacyjną dziecka", edited by Barbara Bokus and Maciej Haman (1992), was issued. It was a ground-breaking publication in Poland, in the sense that it contained the very first studies in Poland referring to the concept of children's Theories of Mind (ToM). Barbara Bokus, in her chapter, reported her studies of children's dyadic interactions, in which she analyzed, among else, how preschool children adjust their communicaton to partner's knowledge, and, in her next book, how they refer to mental states of the characters of narratives, while in my chapter, the review of the state-of-the-arts in the research on children's theories of mind was presented.

Address for correspondence: Maciej Haman, Faculty of Psychology, University of Warsaw, Stawki 5/7, 00-183, Warszawa, Poland. E-mail: maciej.haman@psych.uw.edu.pl

That was the early stage of the world-wide contemporary studies on ToM, and all seemed to be simple then: passing "false-belief" task (FBT) around the age of four was proposed as a corner-stone of ToM (Wimmer & Perner, 1983), a few prototypical versions of FBT were designed and tested in a growing body of research, the interrelations between ToM and some developmental processes studied before (decentration, perspective taking, appearance-reality distinction; Flavell, 1992, 1993) seemed to be well determined, and the sub-population lacking ToM was identified (autistic patients; Baron-Cohen, Leslie, & Frith, 1985, Baron-Cohen, 1989). Unfortunately, since then things became more complicated. First of all, Onishi and Baillargeon (2005) documented passing non-verbal FBT by 15-month-old infants. On the other hand, intriguing irregularities in the ToM task performance were demonstrated in school-age children, adults, and some clinical populations (Henry, Phillips, Ruffman, & Bailey, 2013; Kuhn, 2000; Peterson, Slaughter, Moore, & Wellman, 2016). Last, but not least, widespread use of functional brain imaging methods radically changed contemporary studies of cognition and cognitive development. That, among else, causes that after 25 years, a new critical state-of-the-arts review of ToM research has to be written, discussing such tempting questions as: (a) "Is ToM development continuous from infancy to late childhood, or is there radical conceptual change around the age of four?", (b) "Is there substantial development after passing FBT?", (c) "To what degree does language, executive functions, social interactions, etc., contribute to the development of ToM?", (d) "What areas of children's cognition or behavior (e.g., moral reasoning, social action, etc.) are dependent on ToM development?", (e) "What are neural substrates of ToM and social cognition?"

The size of the current article is too restricted to comprise the whole scope of the live issues in the contemporary ToM research. There are also other review papers which may fill this gap. What I am going to do instead is to focus on two topics: firstly, I am going to point to the contemporary controversies about continuity of ToM from infancy to later phases of development, which arose in the developmental science after publication of the first reports of early understanding of false-belief, and secondly, I am going to touch on the issue of the possible contribution of neuroimaging to the ToM research, and particularly to the continuity debate.

False-Belief Understanding in Infancy and the Issue of developmental Continuity

The history of false-belief test (FBT) began together with the interest in ToM in 1978, when Premack and Woodruff published the article "Does the chimpanzee have a theory of mind?" in the first issue of Behavioral and Brain Sciences. Premack and Woodruff argued, among else, that chimpanzees are able to intentionally deceive other individuals, which requires representing their actual and induced beliefs as a basis of action. This argument was criticized in the commentary by Dennett (1978), who proposed the thought experiment demonstrating that only predicting others action from false-belief provides reliable test for “theory of mind” – understanding that invisible mental states guide actions. In this vein, Wimmer and Perner (1983) designed the false-belief task to be used with preschoolers and found that some 4-year-olds, and majority of older children, but not 3-year-olds, succeed in this task.

Varieties of false-belief tasks have been designed after Wimmer and Perner’s (1983) study, but most of them belong to two general categories: false location (change of location) or false identity (change of content). In the first case, the protagonist observes object X being hidden in the location Y. Then, the object changes its location out of the protagonist’s perceptual access. Finally, in the test phase, the child is asked where the protagonist would search for the object, or where the protagonist thinks the object is. False identity or “change of content” version typically starts with an object hidden in the container suggesting another content (e.g., pencils in the sealed box of chocolate pills). Children are asked what they think is there in the container. Then, the protagonist (one of the experimenters or a puppet) leaves the scene, and the real content of the container is revealed to the child. The test question is what the protagonist thinks is in the box. Additional memory and control questions are asked in both types of the test to ascertain that children correctly understood and memorized all information needed to solve the task. Several versions of these tasks were used in the myriads of studies, which confirmed Wimmer and Perrner’s results, reporting only some minor cultural variability (Lillard, 1998; Shahaeian, Peterson, Slaughter, & Wellman, 2011). That led to the wide agreement between developmentalists that older preschoolers may possess fully-fledged concept of belief, as internal state of mind, independent of reality, linked to other epistemic states, and guiding behavior. Thus, the most basic, adult-like ToM, emerges around the child’s fourth birthday, even if it still develops life-span. This conviction was additionally confirmed by the developmental co-occurrence of the success in FBT and other cognitive tasks requiring appreciation of others’ mental states and understanding partial independence of the mental state content from reality, such as appearance–reality distinction or “level 2” perspective taking (Flavell, 1992, 1993).

Moreover, success in the FBT at least moderately correlated with social abilities such as role taking, joint planning, verbal communication, controlling emotions, moral reasoning, or more general measures of socialisation (Hughes & Leekam, 2004; Jenkins & Astington, 2000; Peterson, Slaughter, & Paynter, 2007). Conversely, patients with autism spectrum disorder (ASD) show severe impairment of social skills and at least delayed or even destroyed false-belief reasoning (see Peterson et al., 2007, 2016). All that seemed to validate FBT as the reliable test measuring one of the most crucial aspects of social cognition, highly relevant both in social perception and action/interaction/communication.

Some doubts arose in early 1990s. First, Zaitchik (1990) suggested that the problem may be not in representing others' beliefs, but rather with the general concept of mental representation – younger children may hold the notion of representation as a copy of reality. In her study, she demonstrated that 3-year-olds expect that the picture from a Polaroid camera would represent current state of reality rather than the state at the moment of taking the picture. Children were indeed able to switch their expectations once they had been instructed how the camera works. Later, Clements and Perner (1994) reported that 3½-year-old children fail in FBT, but look first at the location anticipated from the protagonist's false-belief, thus showing some implicit understanding of false-belief. No such anticipatory looking was observed in 3-year-olds. None of these studies, however, is demolishing for the conception of the child developing his/her fully-fledged ToM around the age of four. "Copy theory of representation" may be a consequence of lacking full ToM, and anticipatory looking at the false-belief indicated location a few months before success in FBT may reflect the necessary developmental transition period.

Another question concerns the pre-requisites for the theory of mind. This question motivated large amount of research done during last 25 years or so, and several possibilities have been proposed and tested. Some form of linguistic/pragmatic determinism of false-belief reasoning was one of the most prominent among them. It ranged from the most radical, pragmatics-oriented version by Katherine Nelson (see Nelson, 1998, Nelson, Plesa, & Henseler, 1998), through Jannett Astington's (1993) view, which stresses the role of general linguistic factors (including syntax and semantics), as well as narrative-communicative competences and social experiences; to grammar-oriented approach of Jill and Peter de Villiers, stressing the role of complement clauses grammar (de Villiers, & de Villiers, 2000; de Villiers, & Pyers, 2002). One of the most persuasive argument comes from the studies of congenitally deaf children (Schick, de Villers, de Villers, & Hoffmeister, 2007). While deaf, non-signing children of hearing parents are highly delayed in false-belief test and impaired in social interaction, those who started to sign as toddlers, develop ToM and social interaction skills at the same time as hearing children, or are only a little delayed (see also Peterson et al., 2016).

Other researchers stressed the role of executive functions (EF) (Moses, 2001). Converging literature reporting behavioral and neuroscientific studies point to the period around the child's fourth birthday as a critical moment in EF development, as well as to the bi-directional relations between EF and social cognition (Zelazo, Carlson, & Kesek, 2008). Executive functions provide satisfactory explanation for Clements and Perner's (1994) results: 3-year-olds get some idea of others' mental states, which correctly directs their attention to the location falsely represented by the protagonist. However, when asked to explicitly decide between false and true location, they are unable to inhibit more prevalent and more recently acquired true location. Recently, Benson, Sabbagh, Carlson and Zelazo (2013) showed that 3½-year-olds who initially perform poorly in classical FBT benefited from ToM training depending on their executive functioning performance. Devine & Hughes (2014) presented the meta-analysis of the large body of studies of EF and FBT performance interrelations, revealing complex, bi-directional, but asymmetrical (especially in the earlier developmental period) pattern. Note, however, that stressing the role of EF in FBT performance may led to the question if younger children's problems with false-beliefs are of conceptual or simply executive nature. In their theoretical analysis of the false-belief task, Bloom and German (2000) raised this objection, and it's gravity also got some empirical support recently (cf. Birch & Bloom, 2003, 2007 – the "curse of knowledge" phenomenon in children and adults; Wang & Leslie, 2016; also executive-pragmatic account by Rubio-Fernández & Geurts, 2013).

Some theories search for the roots of ToM in early, non-conceptual social-cognitive mechanisms developing in infancy. Even if acquired late in the preschool age, ToM may originate from some very early mechanisms of social attention and interaction. Tomasello (1995) stressed the role of gaze following and joint attention (cf. also Tomasello, 2018, for the most recent version of the theory). Meltzoff (1993) pointed to the role of early imitation, which requires mapping between own and other's action. Baron-Cohen and collaborators (see Charman et al., 1997) added empathy to this list. Premack (1990) proposed self-propelled motion as the perceptual basis for mental states ascription, and Gergely, Csibra and collaborators (Gergely, Nádasdy, Csibra, & Bíró, 1995; Bíró, Csibra, Koós, & Gergely, 1998) have shown that infants are sensitive to rationality of action of self-propelled objects even in lack of any other animacy cues. In their view, causal mentalistic action explanations are constructed as theoretical extensions of such early teleological stance (Csibra & Gergely,1998).

Wellman and Wooley (1990) proposed that ToM development progresses from simple desire understanding, present even in toddlers, to more advanced belief-desire scheme, which becomes functional around the age of four. This progression was later extended onto other processes of inferring others' mental states, which allowed to scale the early development of ToM (Wellman & Liu, 2004, Wellman, Fang, & Peterson, 2011; Wellman, 2014)

Implicit understanding of false beliefs before preschool age

There is, however, another possibility which may be levelling for the view of ToM emerging around the age of four, and which, bafflingly, turned fully conceivable after the publication of the study by Onishi and Baillargeon

(2005). Authors of this study constructed a simplified version of the classical "change of location" false-belief task and used it for testing 15-month-old infants in violation-of-expectation (VoE) looking time paradigm. Participants were surprised, and looked longer at the event, when the actor reached for the box which contained the desired object while the object changed location when actor was not present (false-belief condition). In the true belief condition, which was analogous except that the actor was present at the scene all the time, children looked longer at the event in which the actor reached for the empty box, to which the object was moved previously[1]. Although some alternative interpretations are still possible, this pattern is consistent with the hypothesis that infants are able to attribute false-beliefs to other people and to infer expectations about others' behavior from their beliefs, also if conflicting with reality.

Another milestone in studying early understanding of beliefs was the publication by Agnes Kovacs and collaborators (Kovacs, Teglas, & Endress, 2010). The experimental design proposed by Kovacs and collaborators was innovative, and significantly differed from that from Onishi and Baillargeon's (2005) study, or classical FB tasks. Participants were shown short videos in which a ball appeared at the stage, and then was occluded with an opaque screen. A Smurf character was standing next to the stage. Then, the ball left the position behind the occluder, which was visible both for the participant and for the Smurf, and the Smurf was removed. The ball was performing a sequence of displacements returning and leaving the stage, after which the Smurf was returned. This resulted with one of four constellations of beliefs about the final state: P+A+ (both participant and actor–the Smurf–believe that the ball is behind the screen), P-A- (none of the observers believe the ball is behind the occluder), P+A-, and P-A+ (only participant or only actor believes the ball is there). In the final event of the video, the screen was released and revealed either presence or absence of the ball, which might be either consistent or inconsistent with actor's and participant's expectations. The adult participants were instructed to press a key as quick as possible when the ball was present, thus, the Smurf's belief was totally task-irrelevant. Not surprisingly, the reaction times were longer in these trials when the participant did not expect the ball to be present at the release of the screen. However, the inferred Smurf's beliefs also significantly affected reaction times, making them similar (Experiment 1) or even longer (Experiment 2, in which the Smurf did not witness the release of the screen) in P+A- trials than in P-A+ trials. Moreover, reaction times were significantly longer in P-A- trials than in any other condition, in which at least one of the agents (participant or the Smurf)

[1] In fact, two versions of both false-belief and true-belief conditions were designed in this study, which differed in spatial complexity of the object transfer, to minimize the impact of this variable, typically correlating with the experimental condition.

expected the ball to be present. These differences disappeared in the control experiment (the Smurf replaced with a pile of boxes during entire clip), in which only participant's beliefs affected reaction times.

Infants passively observed the same sequences of events, so their looking times, rather than reaction times, were used as indices of violated expectations. In the main trial (Experiment 5), infants looked longer at the scene with no ball when the Smurf expected it to be there in contrast to the child's belief, than at the scene where neither the child nor the Smurf expected the ball. No difference in looking time was found in these two scenes, when no outcome was shown (Experiment 6). Such results are consistent with the hypothesis that both 7-month-old infants and typical adults spontaneously represents not only their own perspective and beliefs, but also perspective and beliefs of other actors involved in the event.

In one condition of the Kovacs et al.'s study (Experiment 2 with adults and Experiment 7 with infants), a pile of boxes replaced the Smurf only in the final phase (during screen release). This manipulation allowed telling at which phase the actor's beliefs are inferred. Despite the lack of the actor during the final event, his beliefs still influenced participants' reactions, which suggests that actor's beliefs are tracked "on-line" and coded at the moment of their formation or change (not at the stage of decision).

There are reasons to grant this study with special status. Firstly, the study was designed particularly rigorously, with several control experiments verifying reliability of the main findings. Secondly, infants participating in this study were only 7-months-old. Thirdly, in this study analogous processing of others' beliefs were shown both in young infants and in adults, providing not only some kind of cross-checking, but also strong argument for continuity in the development of at least some form of ToM abilities. Fourthly, the study demonstrated that processing of beliefs is mandatory and automatic (probably even in young infants). False-beliefs were, in this case, only "by-product" of the event, not its main plot. And finally, the results survived majority of the replication attempts (Kulke & Rakoczy, 2018; but see Kulke, von Duhn, Schneider, & Rakoczy, 2018; Phillips et al., 2015[2] for failed replication attempts).

During over thirteen years after Onishi and Baillargeon's (2005) publication, more than thirty other studies were reported demonstrating early appreciation of false-belief. I am not going to review all of them

---

[2] In the first group of experiments, in a series of 12 studies, Phillips et al. (2015) replicated findings of Kovacs et al. (2010) study, concerning automatic belief tracking in adults. Then, however, they showed that it could be an artefact related to the test events timing. Note, however, that (1) this criticism apply only to the RT measure, so the replications using other measures such as neuroimaging are unaffected by Phillips et al. arguments, and (2) specifically, this argument does not apply to the infants part of the Kovacs' et al.'s study. In their replication of Kovacs et al.'s results with adults, Nijhof, Brass, Bardi, & Wiersema (2016) carefully analysed Phillips et al.'s arguments and found them unlikely in the light of their results. Recently, Kaddouri et al (2019) have shown empirically implausibility of Phillips et al. (2015) argument.

here – there are several comprehensive reviews and discussion papers (see Scott & Baillargeon, 2017; Slaughter, 2015; Wellman, 2017 for critical methodological analysis of over 20 such studies; cf. also Heyes, 2014a). Instead, I am going to focus on some issues following the imposing question if these studies imply developmental continuity from implicit to explicit false-belief appreciation.

Does Early Belief-Tracking Form Continuous Developmental Track with FBT Success at the Age of Four?

Two extreme positions may be taken concerning continuity of development of belief understanding. According to radical discontinuity thesis, any manifestations of early belief understanding are non-conceptual, automated reactions to low level contextual cues and are totally independent of later developed fully-fledged ToM, which takes its representational power from language (semantic and syntactic structures), adopts domain-general flexibility provided by EF development, and is elaborated in social interaction (including communicative behavior). Extreme continuity position may assume that basic concept of belief is ready very early on, and later development does not change conceptual knowledge itself, but relieves executive constraints, adds linguistic and pragmatic means to communicate about beliefs, and, thanks to social experiences, broadens scope of situations in which this knowledge may be effectively used. In reality, however, no one takes any of such extreme position, and actual theories locate somewhere between them. It seems, however, that there are two crucial aspects of continuity debate, which often go together: conceptual continuity, i.e., the issue of conceptual nature of early belief tracking abilities (cf. Carruthers, 2017; Southgate, 2018), and the issue of developmental path continuity, i.e., are later-developed explicit ToM abilities build on the basis of early belief tracking, or are they separate (although somehow linked) developmental paths (this issue is nicely illustrated in Low, Apperly, Butterfil & Rakoczy, 2016). None of these two issues may be easily resolved.

Asking about the conceptual nature of early tracking and reasoning about beliefs, we should look at the methodology of non-verbal, implicit tests of false-belief understanding. Onishi and Baillargeon (2005), and Kovacs et al. (2010), as well as several other researchers, used typical looking time measure in violation-of-expectations (VoE) paradigm, which assumes that infants pay more attention to unexpected event. This measure does not allow determining how deeply the child processed the event. The same issue refers to another attentional measure widely used in the studies of implicit ToM – anticipatory looking. While in VoE paradigm attentional reaction to unexpected action is measured, in anticipatory looking method (deployed in Clements & Perner, 1994), attention allocation precedes expected action. This measure is more universal – may be used with young infants, as well as older children and adults. Both methods, however, assess automatic attentional reactions. Even if the experimental design seems to preclude that this reaction is guided by low level perceptual cues (in Onishi and Baillargeon's experiment, as well as in other studies, false and true belief conditions were counterbalanced for superficial properties), it may be argued that the cue for attention is in familiarity and order of events, or their temporal characteristics, rather than in conceptual interpretation of them. This kind of argumentation was presented in detail by Cecilia Heyes (2014a). Of course, Heyes' arguments of familiarity and novelty are also unwarranted, and one of the problems with them is that familiarity/novelty dimensions had to be *post hoc*, defined separately for each experimental design (the same apply to other minimalist accounts, see Scott, 2014). They show, however, that the up-to-date studies based of attentional indices may not provide fully reliable tests for infants' appreciation of others' false-beliefs.

One may also argue that typical violation-of-expectation or anticipatory looking behavior do not necessarily involve conceptual understanding of beliefs, but simple, automated noticing of either some contingencies or incompatibilities between the event-related cues present in the scene, like the actor's gaze and behavior, object presence and motion, and so on (Gallagher & Povinelli, 2012; Heyes 2014a, 2014b). Such cues are, of course, relevant to understand social interaction, and sensitivity to them may play some role in the development of ToM. Gaze detection, tracking other people's attention, and spatial indexing of objects are present at the very early stage of cognitive development and are computed automatically (Hoehl, Wiese, & Striano, 2008; Hood, Willen, & Driver, 1998; Leslie, Xu, Tremoulet, & Scholl, 1998; Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). Deciding between such mechanisms and more advanced ToM in VoE or anticipatory looking paradigms is also unwarranted. Moreover, Heyes (2014b) argues that such "submentalizing" procedures may be sufficient even for adults to effectively interact in standard social situations. However, even if we pass over it and insist on granting young children with some more advanced ability to track content of others' mental states (including false-beliefs), it may still be highly contextually restricted to on-line inferring of other's object and object-location knowledge from perceptual access, in order to predict allocation of the actor's attention or object-directed action only (Apperly & Butterfill, 2009, proposed a term, "belief-like states", to refer to such attention-object-action relations). While such ability may still be one of the distal precursors of true conceptual understanding of beliefs, it is far insufficient to credit infants and toddlers with fully-fledged conceptual ToM. Meert, Wang, & Samson (2017), however, controlled their experimental materials for protagonist gaze-object associations and, despite that, found a robust effect of belief tracking in adults.

There are, however, some other procedures used to demonstrate implicit false-belief processing in infants and toddlers which are not attention-based, and are partly resistant to criticism presented above. Buttelmann, Carpenter and Tomasello (2009) demonstrated that eighteen-month-old children infants are able to infer actor's goal from actor's belief and adapt their own helping behavior to this goal. In the false-belief condition, an actor placed a toy in one box and left the room. When the actor was out, the experimenter showed the child how to lock and unlock the box with a pin, moved the toy to another box and locked the boxes. Then, the actor returned and unsuccessfully tried to open the box (now empty) in which he/she put the toy in. At this moment, the children were allowed to approach the boxes and encouraged to help the actor. More than 70% of eighteen-month-olds and 80% of sixteen-month-olds unlocked the box containing the toy, although the actor's behavior was directed to the empty box. As far, this is not surprising – the child's behavior might simply reflect his/her own knowledge about the object location. However, in the true belief condition in which the toy was transferred to another box in presence of the actor, more than 80% of 18-month-olds helped the actor to unlock the empty box. That was not so clear with 16-month-olds, as almost half of them opened the box with the toy, but the proportion was still below that found in the false-belief condition. Although this procedure is also based on tracking object location, it shows that the consequences of very young children's belief reasoning are more far-reaching (helping action), and thus involvement of some kind of conceptualization is more likely here.

Moreover, although most of the studies used change of location paradigm, which may be highly susceptible to tracking gaze of object location cues, there are also studies that showed infants' ability to deal with false-belief in change of content task. For example, Buttelmann, Over, Carpenter and Tomasello (2014) also adapted helping behavior as a dependent measure, but the task was to help an actor by providing some blocs or sponge, depending on what the actor believes is in the box. This paradigm seems to be especially resistant to the critics concerning low-level contingencies, gaze and object location tracking or object properties, as all these factors remains constant across all experimental condition, and the only difference lays in the agent's belief, which should be inferred by the child being tested.

Indeed, an argument still could be made that both Buttelmann et al.'s (2009, 2014) studies measured child's object-directed action, and thus added only a little of flexibility to the "perception → (knowledge) → object-directed attention → object-directed action" scheme. It is not clear if helping behavior requires understanding beliefs or, rather, other conceptualization, e.g., teleological understanding of action (which is also ToM-related, but not constituent for it), is enough here (Priewasser, Rafetseder, Gargitter, & Perner, 2017). The very recent study by Stojnić and Leslie (2018) documented that the use of the results of belief tracking are not restricted to inferring goals and object-directed action plans. In this study, an actor called the dog by its proper name – Fido, put it to the box and left the scene. Then, the dog jumped out the box and another identical dog jumped in. The actor returned with another actor and, pointing to the box, said, “There is Fido in the box.” Then, both actors left the scene again, the first dog returned and both dogs were presented side by side. Fifteen out of 20 3-year-olds pointed to the correct dog when asked to point at Fido. This study extended previous results of Scott & Baillargeon (2009) with 16- and 18-month-olds in which individual identity, rather than presence or absence of the object, was the content of the belief (two identical penguins changed location when the actor was not attending). Similar results were found in 2-year-olds in preferential looking paradigm in Stojnić’s (2017) study (but see Oktay-Gür, Schulz, & Rakoczy, 2018 for arguments against full conceptual understanding of beliefs about identity in young children).

Finally, Wang and Leslie (2016) used anticipatory looking in their study of 2- to 3-year-old children’s and adults’ implicit false-belief tracking. They varied, however, executive load needed to solve the conflict between the agent’s false-belief and the participant’s true knowledge about reality. They revealed that gaze-allocation between false and true locations is highly dependent on the conflict load in both adults and young children (which is similar to “curse of knowledge” phenomenon in explicit false-belief processing; cf. Birch & Bloom, 2003; 2007). Wang and Leslie’s results directly speak against low-level contingencies hypothesis, proposed by Heyes (2014a), and show sensitivity of belief processing to executive load related to conceptual conflict not only in adults, but also in children below the critical age of four. Authors also discuss other arguments for the rejection of early ToM competence hypothesis, and conclude that contrary to these arguments, early abilities to represent others’ mental states should be considered as conceptual “theory-like” structure. Careful argumentation for conceptual nature of belief processing in children much before succeeding in explicit FBT is also provided by Southgate (2018).

Some interesting arguments for conceptual continuity of ToM have been proposed by Tomasello (2018). Like in his previous account of social cognition (cf. Tomasello, 1995), this author seeks for sources of unique human social-cognitive abilities in shared intentionality, and behavioral and mental coordination with the other person, which requires joint attention, linguistic communication, and advanced executive functions. Uder Tomasello's account, the problem lies not in the basic conceptual constructs, which may be shared even with some non-human species, but in coordination of different perspectives represented with the use of these concepts (which leads to more advanced ToM). Paradoxically, 3-year-olds’ troubles with both explicit and some versions of implicit tasks, regarded as an argument for distinct paths of development of explicit and implicit ToM abilities, may in fact represent

U-shaped pattern of age-related performance in FBT. Maybe, 3-year-olds start to develop abilities to represent not only subjective (their own and other person's) perspectives, but also objective perspective, and use this concept too widely, assuming that people typically behave according to objective state of reality.

Two main arguments were provided in the debate on developmental path continuity. The first one concerns correlation (or lack thereof) between explicit and implicit false-belief reasoning. The second one concerns disunity/unity of ToM abilities before and after passing explicit FBT. Some studies, both longitudinal and cross-sectional, provide mixed evidence for the correlation between performance in spontaneous belief-tracking and verbal false-belief task in adults and children around the critical age of four (see Grosse Wiesmann, Friederici, Singer, & Steinbeis, 2017), although most of them report at least some positive interrelations (Low, 2010; Thoermer, Sodian, Vuori, Perst, & Kristen, 2012; Sodian, 2016; Oktay-Gür et al., 2018). Authors of some of the abovementioned studies point to the lacking correlation as undermining the thesis about continuity from implicit belief tracking in infancy to the success in explicit FBT around four. In my view, however, determining such a correlation raises some serious theoretical and methodological issues. To determine if someone spontaneously tracks some other's belief, we assess allocation of attention, usually looking time or gaze location order. For instance, in VoE paradigm, we expect the child to look longer at an unexpected event, i.e., when actor behaves inconsistently with her/his expected beliefs. It does not imply, however, that proportion of looking time exactly indicates belief tracking skills. Allocation and duration of gaze is a property of executive, not conceptual, system. It may easily happen that a child with better developed executive control would look briefly at the expected outcome, only a little longer at the unexpected outcome, and, then, disengage attention, having just acquired all information needed. On the other hand, the child with poor executive skill would stick to the unexpected outcome, and look at it much longer both in the sense of absolute time and proportion, simply because of poor ability of disengaging attention. Similarly, switching gaze between true and false location may rather reflect conflict processing (as probably was in Wang & Leslie, 2016, study), only indirectly indicating understanding beliefs.

Another argument against correlational approach may be derived from the assumption that understanding of beliefs constitute radical conceptual change. In a consequence, success in classical FBT should be of all-or-none nature. That is what was found in several studies (although, see Baker, Leslie, Gallistel, & Hood, 2016 for convincing demonstration that transition from failing to succeeding in FBT is a smooth developmental process lasting a few months). The crucial expectation in false-belief test is that children before conceptual change would answer test question on the basis of their own beliefs about reality, and not randomly. However, correlational studies typically use a success rate in a series of ToM tasks or, even worse, some classification of children's justifications of their choices in FBT for scaling the subject's advancement in understanding mental states, sometimes with uneven answers counted as "half point". Again, uneven answers or mixed performance in verbal FBT may rather document the problem with conflict management, but not understanding, which should be all-or-none except short transitional period. Correlation between these kinds of measures may then have nothing to do with assessing relations between implicit and explicit belief processing[3]. Assuming that, depending on the task, different executive demands impact measured variables, there is no reason to expect any strong correlation between them.

Second argument used in path continuity debate is similar, and at least partly falls under similar criticism. Some authors (Oktay-Gür et al., 2018; Poulin-Dubois & Yott, 2018; Yott & Poulin-Dubois, 2016; see also Low et al., 2016) point to a lack of generalizability between different implicit ToM tasks (change of location, identity, perspective taking) and methods (VoE, anticipatory looking, helping behavior). Children (and sometimes adults) who pass one task may not pass another task in another trial. On the other hand, congruency in different explicit ToM tasks performance was reported even in the earliest studies (see introductory part of this paper) and replicates in contemporary research. But again, although researchers try to make their experimental designs maximally equalized, as far as we use implicit measures, different tasks and contexts pose different attentional, executive, and memory requirements. Executive and linguistic demands of explicit ToM tasks are more unified (or at least scalable, which allows better control over them).

Nonetheless, lack of well-established correlation between implicit and explicit beliefs processing and lack of consistency in implicit ToM performance are often used as arguments against continuity between implicit and explicit belief reasoning. Basing on them, several researchers adhere to the two-system approach proposed by Apperly and Butterfill, 2009; Apperly, 2013; Low, Apperly, Butterfill and Rakoczy, 2016). According to this view, "System 1", which is evolutionary ancient, automatically tracks "belief-like" states using direct cues, like agent's gaze. "Belief-like" states are relational

[3] Note, however, that this criticism may not apply to continuous measures assessing sequences of developmental achievements in ToM (see Wellman & Liu, 2004) in which different concepts constituent for ToM are assessed. There is also an intriguing modification of change of location FBT, proposed by Sommerville, Bernstein and Melzoff (2013) in which participants are asked where the actor would search for the moved object in a continuous space (sandbox), and the distance of the searched location from the location known by the participant as true and from the location suggested by the actor's false-belief, is claimed to indicate "level" of appreciation of other's beliefs. In this procedure, participants were deliberately distracted between the presentation of the cover story and the test, to prevent them from relying on direct perceptual cues. This task correlates with standard FBT, and was proven to be useful in a wide scope of studies with very young children, as well as adults, elderly, and atypical populations, and in neuroimaging studies. However, in light of the previous discussion, it seems more likely that this task provides good measure of conflict management abilities, which are indispensable part of false-belief processing, rather than conceptual understanding of beliefs.

attitudes, in which content may be distinguished by the relations between the agent, object, and its location or other properties. “System 1” is efficient, as it does not load executive and linguistic resources. It is, however, inflexible – it applies to the typical actor-object-location relations only. On the other hand, “System 2” represents beliefs in canonical form as propositional attitudes, and distinguishes between their contents by conforming truth conditions and considering their aspectuality. “System 2” extensively engages executive resources and is recruited by tasks that require declarative expressions of, or deliberation about, beliefs, and does not require direct situational cues like actors’ gaze direction (Low et al., 2016).

Proponents of the two-system accounts stress the developmental autonomy of “System 2” as an advanced conceptual construction, achieved thanks to linguistic support, appropriate social experiences, and developments in cognitive control (executive functions), although they do not preclude some secondary ontogenetic and functional dependences of the “System 2” on the “System 1”. Detailed discussion of arguments in favor of two-system (also beyond weak consistency of performance in spontaneous and deliberate ToM reasoning) forming two separate developmental paths may be found in Low et al., (2016), so I skip presenting them here. There is, however, one argument which comes from autism research and seems to favor independent paths for development of explicit and implicit ToM abilities. Some high-functioning patients diagnosed with autism spectrum disorder (ASD) are severely impaired in automatic belief tracking. They develop, however, relatively good skills of explicit belief reasoning in standard false-belief tasks (and other task requiring higher-order ToM reasoning, cf. Scheeren, de Rosnay, Koot, & Begeer, 2013).

This argument may, however, be kind of a “double-edged sword”, since it may be partly demolishing to the concept of explicit ToM as a cognitive basis of social skills (which is widely accepted, cf. Hughes & Leekam, 2004; Moran et al., 2011; Scheeren et al., 2013). Despite their relatively good performance in false-belief task, high-functioning ASD patients’ social skills remain highly restricted (Moran et al., 2011; Peterson, Slaughter, & Paynter, 2007; Peterson et al., 2016). Thus, if verbal FBT success is assumed to be diagnostic for ToM, explicit ToM may not be a precondition of socialisation – ASD patients may possess explicit ToM but are unable to build standard social relations. At the same time, the relations between implicit ToM reasoning and social skills were not posed under question in up-to-date research (see Schuwerk, Vuori, & Sodian, 2015). If so, do we really need explicit ToM in social cognition and behavior? Fortunately, we are not forced to reject ToM as a cognitive basis of social behavior. It may be argued that explicit ToM abilities in high-functioning ASD patients are developed as some compensations, which emerge along atypical path (see Farrar, Benigno, Tompkins, & Gage, 2017). I find this hypothesis feasible, if so, however, the argument about explicit ToM-like abilities in ASD becomes irrelevant in the main line of the continuity vs. two-system debate.

There are also some other problems with the two-system, and other discontinuity accounts. Firstly, as I have argued above, spontaneous belief tracking is not so inflexible, and subserves relatively large range of behaviors – much larger than expected from “System 1” properties in the two-system view (see, however, Low et al., 2016, for explanation consistent with the two-system theory). Secondly, there is increasing evidence that when executive and linguistic requirements of the task are released, even children below three years of age may sometime pass versions of false-belief task requiring explicit false-belief reasoning. For instance, Rubio-Fernández and Geurts (2013) designed “Duplo task” – a version of classical explicit change-of-location task, which, however, minimizes linguistic, pragmatic, and working memory load, and specifically does not require the child to focus his/her attention on the object and its current location. In this task, even 2½-year-olds succeeded. On the other extreme, there is the “curse of knowledge” phenomenon – even adults rely on their knowledge of reality in false-belief reasoning in some pragmatically misleading or executive-demanding situations (e.g., in case when the correct choice is objectively implausible; Birch & Bloom, 2007; Wang & Leslie, 2016). Interesting version of the pragmatic-executive constraints hypothesis has been proposed by Helming, Strickland and Jacob (2014, 2016). According to these authors, most of the explicit false-belief tasks require complex pragmatic turn from the child: from third-person observer of the actor’s false-belief perspective to second-person perspective in interaction with experimenter, who is asking test questions about the actor’s belief, with a necessity to control child's own first-person perspective. Complexity of establishing equivalences of these perspectives, rather than managing belief-reality conflict (which may be a part of implicit belief tracking as well), is the main factor causing failure in FBT before the age of four. Thirdly, not all attributes of “fine-grained” belief concept, claimed to be part of “System 2”, seem to be “at place” just at the moment of supposed emergence of “System 2”-based ToM around the age of four. Some studies show, for example, that understanding of aspectuality of beliefs emerges about a year later than the ability to pass standard false-belief task. Even in tasks carefully controlled for linguistic and executive demands, only about 50% of 5-year-olds pass double identity task while majority of them pass unexpected content false-belief task, fully equivalent in complexity and executive demands (as well as change of location FBT; Gut, Haman, & Gorbaniuk, under review). And finally, there are a few studies whose results suggest that spontaneous abilities to track “belief-like” states also partly rely on communicative experiences and executive control (Schneider, Lam, Bayliss, & Dux, 2012; Wang & Leslie, 2016), although they are less demanding than explicit belief reasoning, which provide indirect support for the thesis about building advanced ToM on the basis of earlier belief-tracking abilities. Indeed, Low et al., (2016), argue that some components of "System 1" may pose demands on the cognitive resources. However, accepting this claim makes two-system theory ill determined. The abovementioned arguments show the state-of-the-arts more consistent with the continuity hypothesis (although do not corroborate it directly), and troublesome for the two-system account. In this light, one may argue that evidence referred to by the two-system theory proponents may be more parsimoniously accounted for in the model proposed by Carruthers (2017), in which single set of concepts may be either deployed automatically in response to some salient contextual cues, or more deliberately and flexibly in correspondence to domain-general requests.

There is indeed one important caveat concerning early ToM competence which should also be mentioned here. Although more than thirty studies involving infants or toddlers, and reasonable number of studies with adults, reporting positive results indicating automatic processing of beliefs over lifespan have been published, often replicating and extending previous findings, increasing number of systematic replication attempts gave mixed or negative results (cf. Kulke & Rakoczy, 2018; Kulke, Reiß, et al., 2018; Kulke, von Duhn, et al., 2018; Phillips et al., 2015; Poulin-Dubois & Yott, 2018; see also other papers in the special issue of Cognitive Development, Vol. 46, pp. 1-124). Even if part of the evidence seems persuasive, effects under study may be fragile, and depend on factors which may not yet be known (but see Baillargeon, Buttelmann, & Southgate, 2018 for a careful analysis of possible causes of replication failures). In this light, available evidence may be taken in favor of either continuity or discontinuity thesis, depending on which aspects of it would be stressed. That recommends caution in interpreting data collected up to day. For these reasons, some alternative methodological approaches may be desirable, and the next section of this article will be devoted to them.

Summing-up, young infants as well as older children and adults spontaneously track others' beliefs (at least those on-line acquired by perceptual access to object's identity and location), and use them to predict states of reality, actor's attention allocation, and object-directed actions, to infer actor's goals and intentions, including representational intentions (naming), as well as to direct subjects' own attention, and to plan their own object-directed and social (actor-directed) actions. In opposition to some critics, this ability seems to be flexible enough to be considered as some form of conceptual knowledge. But the question remains, how this knowledge is related to the ability to solve false-belief task in its explicit, verbal form which typically emerges only around the age of four years. Some arguments for conceptual break-out around the age of four are not so easy to abolish. Higher-order ToM conceptual abilities seems to be clearly linked to explicit false-belief performance (usually develops few years after succeeding in explicit false-belief task), when no clear relation between them and implicit belief tracking was established. It seems that ability to explicitly overcome working memory and executive restrictions, together with rejecting copy of reality, true-belief, or own-belief-based hypotheses, opens doors to more advanced conceptual constructions. Language (appropriate syntactic constructions and mental vocabulary), and some socialisation training may also extensively contribute to this process. However, such a view is at least equally consistent with two-system theory as well as continuity theory. It seems that the debate on continuity vs two-system accounts of ToM development has to be moved onto a different level.

Neural Underpinnings of ToM and the Continuity Debate

ToM was one of the first area of cognition investigated with use of contemporary methods of functional brain imaging. The earliest studies, typically utilizing positron tomography (PET) technology, gradually replaced with fMRI, revealed several brain sites engaged in ToM task processing, among which medial prefrontal cortex (mPFC) seemed to play a central role (see Fletcher et al., 1995, and for the review, Gallagher & Frith, 2003). Subsequent studies, mostly utilizing fMRI (or, in some cases, EEG and TMS), have led to re-model ToM and social cognition networks in the brain. In summary (see meta-analyses by Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Schurz, Tholen, Perner, Mars, & Sallet, 2017; Schaafsma, Pfaff, Spunt, & Adolphs, 2015), processing socially-relevant information (including processing others' internal epistemic states, inferring intentions from motion and from eyes, reading emotions, etc.) engages several subnetworks with local hubs within bilateral temporal, parietal, and lateral and medial prefrontal cortices, as well as in subcortical brain areas, with central, integrative role played, however, by temporo-parietal junction (TPJ), especially in the right hemisphere. Right TPJ is also specifically engaged in belief processing, although anterior medial prefrontal cortex, temporal poles, and precuneus seems to be involved too. Interestingly, these networks converge also with so called "default mode network" – pattern of functional cortical connectivity observed during rest, and related to mind-wondering and feeling of self (Mars et al., 2012; Spreng & Grady, 2010).

Finding overlap between main sites for belief processing and integration of socially-relevant knowledge may provide additional support for the central role of the ToM in social cognition, however, for the purposes of the current analysis, more important is if implicit and explicit belief processing share the same neural network. Only a few studies touched this issue directly. In the earliest one, Schneider, Sloughter, Becker and Dux (2014) collected fMRI scans from participants passively watching "Sally and Anne" false-belief clips. Although eye tracking data documented typical anticipatory looking pattern of false-belief tracking, only precuneus and anterior sites in the left parietal lobe exhibited stronger activation contrast between false and true belief conditions. Crucially, no such contrast was found in rTPJ, which is characteristic for explicit false-belief processing. Importantly, however, this structure showed significantly increased activation in both false and true belief conditions. Later studies did not confirm this dissociation between implicit – explicit belief processing. Bardi et al. (2016) used a task modelled after Kovacs et al.'s (2010) study in two versions: implicit and explicit. In both versions, participants' task was to press the reaction key when the ball was present in the outcome (after the screen release), and in some trials a question was asked after the trial. In the implicit condition, the question was "Did Buzz [the observer in the clip] have a blue cap?" and in explicit condition, the question was "Did Buzz think the ball was behind the screen?" For both questions, 50% of correct answers was "yes" and the remaining 50% was "no". Participants were performing the tasks in a MRI scanner in two sessions, the implicit condition always first. After the implicit condition session, participants filled in a debriefing form to check if they were aware of belief manipulation. The study replicated behavioral results from Kovacs et al. (2010). Brain activations were analyzed in two time windows: belief formation (Buzz watches the ball moving behind or out of the screen) and outcome phase (the screen is released). During belief formation, increased activation was found in a couple of sites, including right angular gyrus (located in rTPJ). No effect nor interaction involving the condition (implicit vs. explicit) was found. In the outcome phase, violation of expectation based on observer's belief resulted with activation in anterior medial prefrontal cortex (amPFC), and violation of expectation based on participant's belief activated adjacent part of amPFC. Again, no effect of explicitness was found. Interestingly, when activation in preselected ROI (region of interest), i.e., in the right TPJ was analyzed, specific activation was found only in case of false-belief with positive content (observer falsely believed that there is a ball behind the screen). Such an effect was shown for the first time in the study by Kovacs et al. (2014) in automatic belief tracking, in which only implicit task was used, and if this effect would not replicate in explicit belief task, it might provide an argument for restricted belief representation in spontaneous belief tracking. Indeed, the same positive belief constrain affected rTPJ activation in explicit belief condition. In a more recent study, Nijhof, Bardi, Brass & Wiesema (2018) tested also autistic adults in the same paradigm. While at the behavioral level no difference was found in explicit false-belief processing between ASD and control groups, no increased activation in rTPJ was observed for false-belief contrasted to true belief condition in ASD group. This result supports the hypothesis that high-functioning ASD patients' success in explicit FBT has different bases than in healthy subjects. Bardi, Six and Brass (2017) have also shown that rTMS (repetitive transcranial magnetic stimulation), which temporally blocks neuronal activity at the stimulated location, applied over rTPJ, impairs implicit false-belief processing. Similar overlap between brain activation pattern in explicit and implicit false-belief processing was found in Naughtin et al.'s (2017) study, with use of the more emblematic "Sally and Anne" task type.

It is also worth to mention two studies by Hyde and his collaborators. Contrary to the previously reported works, in which MRI was used, Hyde's group utilized near-infrared functional spectroscopy (fNIRS). Like fMRI, fNIRS determines localization of cortical activity estimating changes in cerebral blood oxygenation level (BOLD). NIRS have some disadvantages when compared to fMRI – it is less precise and does not allow to approach deeper cortical or subcortical brain structures (for example large parts of precuneus and ventromedial prefrontal cortex, among those involved in ToM). However, it has at least one advantage critical in developmental research: it is much more patient-friendly. NIRS equipment does not generate noise, preparation for testing does not take a lot of time, the probe may be done in much more natural settings (e.g., the youngest children may sit on their parent's laps), and is less sensitive to patient motion, which not only fit better typical child behavior, but also allows for much richer choice of reaction options. In the first study, Hyde, Aparicio Betancourt and Simon (2015) tested adult participants with implicit version of unexpected change of location FBT, similar to that used by Onishi and Baillargeon (2005). Recordings of fNIRS from a channel covering rTPJ (previously identified as active in explicit false-belief reasoning) revealed increased activity of this structure in all scenes requiring inferring the protagonist's mental states. Moreover, false-belief condition engaged this structure even stronger than other belief conditions. In the next study (Hyde, Simon, Ting, & Nicolaeva, 2018), almost the same procedure was applied to 7-months-old infants and revealed very similar results. Although only one (but crucial) site in ToM brain network was monitored here, these studies show that explicit and implicit belief processing may share at least some neural bases (located in rTPJ), and that functional properties of this site of ToM brain network are established at the very early stage of development.

Last reported study leads to another question which should be asked here: are there any specific changes in the brain structure or function correlated with the developmental success in the explicit false-belief task? As far, this problem was approached in only very few studies. Sabbagh, Bowman, Evraire and Ito (2009) demonstrated that better performance in explicit ToM tasks correlates with changes in current-source density in alpha-bound EEG oscillations in the areas associated with ToM: rTPJ and mPFC, when controlling for executive functions and age. Such differences in alpha-bound activity are supposed to indicate maturation of given brain structures.

Since prefrontal cortex functioning depends mostly on dopaminergic connections, the same research group has demonstrated in subsequent studies that ToM performance in preschoolers is related to the level of dopamine (Lackner, Bowman, & Sabbagh, 2010) and part of individual differences may be linked to one of the genes coding dopamine receptors (Lackner, Sabbagh, Hallinan, Liu, & Holden, 2012). In another study, Grosse Wiesmann, Schreiber, Singer, Steinbeis, & Friederici (2017) used diffusion-weighted MRI imaging with 3- to 4-year-old children. They found that higher false-belief scores correlated with increased functional anisotropy within the right TPJ and pMTG (posterior medial temporal gyrus, which is adjacent to TPJ), right ventral mPFC, right precuneus, and left MTG. Functional anisotropy indicates white matter maturation in these ToM related sites. Connectivity paths between these regions were also found to be stronger in FBT-succeeding children. These correlations were not affected by executive and linguistic performance level All studies presented above provide some support for dependence of the success in false-belief task from the developmental processes in the brain which cannot be reduced to general, age-related, development, nor, more specifically, to executive function network or linguistic development. At the same time, however, they show involvement of the very same brain regions before and after the success in explicit FBT.

Very recent study by Richardson, Lisandrelli, Riobueno-Naylor, & Saxe (2018) made the next step towards the same direction. Large sample of children (N = 122), aged three to twelve years, and adults (N = 33) participated in the study, in which they passively watched an animated video during functional MRI scan. The protagonist of the cartoon (a stork Peck, from Pixar Studio animation, "Partly Cloudy"), experienced a series of bodily sensations (pain in most cases) and mental states, such as beliefs, desires, and emotions. After the fMRI scan, children completed a battery of tasks including executive function probe and explicit false-belief task, which allowed to categorize the participants into three classes: FB-passers, FB-inconsistent, and FB-failers. The analysis of fMRI BOLD signal revealed two separate brain networks, one active when the protagonist experienced pain, and the second one active when internal mental states of the protagonist were to be inferred. This second one consisted of the sites previously documented in ToM imaging studies: temporoparietal structures (right and left TPJ, right precuneus) and medial prefrontal cortex. Both networks were functionally separable even in the youngest group. Importantly, however, their functional specialization increased with age, ranging from no between-network correlation in 3-year-olds, to negative correlation in the older groups, and interconnectivity (especially between temporoparietal and medial prefrontal sites). These results are in line with previously reported studies, and extend the result from Gweon, Dodell-Feder, Bedny, and Saxe's (2012) study, in which increasing specialization of ToM network between the age of 5 to 12 years was shown. What, however, is the most important in the current study is that although no explicit belief tracking was prompted during the clip watching, age-related change in the brain activity pattern parallels that revealed in explicit ToM tasks reported previously. Indeed, no apparent differences in functional responses within ToM network was found between children who failed and those who passed false-belief task in the post-test. All that provides more direct support for the continuity hypothesis and leads to suggestion that the success in explicit FBT depoends on maturation of the ToM network, which however is functional at the basic level early on, and may subserve implicit belief tracking at the earlier stages. Additionally, one result, which was not discussed by authors in detail, may also be of special interest here: in the youngest children, medial-frontal and temporoparietal parts of ToM network seem to be only loosely interconnected. Any attempts to functionally fractionate "social brain" networks (e.g., Schurz et al., 2014) are mostly speculative at this stage. Nevertheless, conflict monitoring is often indicated as a function of anterior cingulate cortex and adjacent regions of amPFC, even outside ToM research (cf. Carter & Veen, 2007). Partial immaturity of this site and its weaker connectivity with TPJ (which, in turn, is sometimes claimed to be responsible for belief tracking), may provide an explanation for the difficulties of younger children with explicit FBT, which may lay somewhere between domain-specific (conceptual) and domain general (executive) resolutions of this problem.

Summing-up, it seems that the same neural network is active during implicit and explicit belief processing from the earliest stage of development. This network, however, undergoes the maturation process, increasing connectivity within specific subnetworks and between them, and increasing functional specialization in the sense of suppression of activity in competitive networks (e.g., pain empathy). This is a continuous process of development, however, significant change was noted in the preschool age, and in some studies, it paralleled explicit false-belief success, even if controlling for age or executive functions. Such data seem to fit the continuity hypothesis better, but still allow to assign a "milestone" status to the moment of passing explicit FBT. Of course, this is not enough to reject the two-system theory and even some more radical discontinuity accounts. To do that, explicit and implicit belief processing should be analyzed with much better resolution (both in the sense of function, see Kampis, Fogt & Kovacs, 2016; and activation pattern analysis, see Jacoby, Bruneau, Koster-Hale, & Saxe, 2016) to check if they do not show disjunctive patterns of activity within the same general brain networks. Importantly, however, current neuroimaging studies overcome some objections raised by Apperly (2013), as they shift from localist, subtractive methodology to network-based analyses.

General Conclusions

Theory of mind in general, and false-belief task in particular, have competed with executive functions for the position of the most studied problem in the contemporary cognitive-developmental science. Enormous amount of data has been collected over 40 years after Premack and Woodruff's seminal paper, 35 years after Wimmer and Perner's first study with FBT, and 25 years after our first Polish publication mentioning children's ToM, and most of this data confirmed central role of ToM in social cognition, communication, and communicative behavior. Perhaps, the most intriguing discovery in the area of ToM was the demonstration of very early (just in the first year of life) ability to track others' beliefs (also false), which guides the child's expectations concerning behavior of other people, as well as the child's own social behavior (e.g., helping). Fully analogical ability of automatic (or semi-automatic), non-volitional belief-tracking was demonstrated also in adults. This discovery, almost obligatorily, opens discussion about the earliest conceptual competences in this domain, being one of the most basic areas of cognition.

The relation between performance in implicit belief tracking and traditional explicit false-belief tasks is not, however, yet determined. Some researchers propose the two-system theory, with only weak connections between implicit and explicit systems (only this second is claimed to to involve conceptual reasonig), while others opt for the continuity hypothesis. While evidence discussed in this paper fits better the continuity thesis, it remains unclear if achieving success in explicit FBT reflects only continuous maturational process in executive and communicative competences, or still constitutes a milestone of important, qualitatively new, stage in the development of social cognition (even if based on increased executive abilities). Belief construct deployed in the spontaneous belief tracking process may be abstract (in the sense of keeping identity across contexts) and conceptual in principle. However, new windows to new conceptual constructions around the age of four may be opened along with better developed executive control and better monitoring of conflict inherently involved in false-belief processing, but also with acquisition of mental lexicon facilitating instantiation of the conceptual components, and providing, together with conditional syntax, productivity and displacement typical to language. These new developments: full understanding of aspectuality, or second-order ToM concepts, sometimes may even be achieved along alternative path, without the ability of spontaneous belief tracking (as in the case of some well-functioning ASD patients).

Neuroimaging tools, available in last thirty years or so, may significantly contribute to this debate. Unfortunately, these tools are typically not children-friendly. Some fMRI studies, imaging brain activity during explicit and implicit processing of others' mental states (including false-beliefs), showed similar pattern of activation in both conditions, but only a few of them involved children around or before their fourth birthday. Nonetheless, neuroimaging studies shed some additional light on the current issue. Particularly, brain connectivity studies show presence of the main ToM network just at the earliest stage of development. However, the strength of the connections, both local and distal, and selectivity of activations within this network increase with age. Importantly, the same network is engaged in both, implicit and explicit belief reasoning. Additional opportunity to study brain activity at the transitional age may be provided by recently developed fNIRS technology. This method is cheaper, easier to apply than fMRI, and, most importantly, children-friendly.

Several other important achievements in the development of the ToM area from the last twenty-five years were not presented here. A lot was learned about other early manifestations of inferring others' intentions, and other mental states from gaze, communicative cues, or motion pattern. For instance, interesting relations between ToM and face processing were also found. We know more about the nature of ToM impairment in ASD, and specific problems also with ToM development and reasoning in other atypically developing populations. In this paper, I have skipped the highly controversial issue of the mirror neurons system and its role in ToM, as well as some very recent, intriguing developments in Premack & Woodruff classical question, "Does the chimpanzee have a theory of mind?" Recently, which may arise as a bombshell, Krupenye, Kano, Hirata, Call and Tomasello, 2016, 2017 (follow-up study) reported in *Science* two experiments showing the ability to anticipate others' behavior from false-belief in chimpanzees. Only a few months later, Buttelmann, Buttelmann, Carpenter, Call, & Tomasello (2017) adapted their helping task to animal studies and got similar results. Subsequently, some abilities to track others' knowledge have been demonstrated in rhesus monkeys (Drayton & Santos, 2018). And finally, a few months ago, Rabinowitz et al. (2018) reported that deep-learning machines can also do that.

Acknowledgments

Preparation of this manuscript was supported by National Science Centre, Poland, grant no. 2017/25/B/HS6/01624.

References

Apperly, I. (2013). Can theory of mind grow up? Mindreading in adults, and its implications for the development and neuroscience of mindreading. In Baron-Cohen, S., Tager-Flusberg, H. & Lombardo, M. (Eds.) *Understanding other minds: Perspectives from developmental social neuroscience* (pp. 72–92). Oxford, UK: Oxford University Press.

Apperly, I. A. & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological review*, *116*(4), 953.

Astington, J. W. (1993). *The child's discovery of the mind*. Cambridge, MA: Harvard University Press.

Baillargeon, R., Buttelmann, D., & Southgate, V. (2018). Invited commentary: Interpreting failed replications of early false-belief findings: Methodological and theoretical considerations. *Cognitive Development*, *46*, 112–124.

Baker, S. T., Leslie, A. M., Gallistel, C. R., & Hood, B. M. (2016). Bayesian change-point analysis reveals developmental change in a classic theory of mind task. *Cognitive psychology, 91*, 124–149.

Bardi, L., Desmet, C., Nijhof, A., Wiersema, J. R., & Brass, M. (2016). Brain activation for spontaneous and explicit false belief tasks overlaps: new fMRI evidence on belief processing and violation of expectation. *Social cognitive and affective neuroscience*, *12*(3), 391–400.

Bardi, L., Six, P., & Brass, M. (2017). Repetitive TMS of the temporo-parietal junction disrupts participant's expectations in a spontaneous Theory of Mind task. *Social cognitive and affective neuroscience*, *12*(11), 1775–1782.

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of child Psychology and Psychiatry*, *30*(2), 285–297.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition*, *21*(1), 37–46.

Benson, J. E., Sabbagh, M. A., Carlson, S. M., & Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Developmental Psychology*, *49*(9), 1615.

Birch, S. A., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science*, *14*(3), 283–286.

Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false-beliefs. *Psychological Science*, *18*(5), 382–386.

Biro, S., Csibra, G., Koos, O, & Gergely, G. (1998). Understanding rational action in infancy. *Psychology of Language and Communication*, *1*(2), 29–38.

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.

Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J., & Tomasello, M. (2017). Great apes distinguish true from false-beliefs in an interactive helping task. *PLoS one*, *12*(4), e0173793.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false-belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337–342.

Buttelmann, D., Over, H., Carpenter, M., & Tomasello, M. (2014). Eighteen-month-olds understand false-beliefs in an unexpected-contents task. *Journal of Experimental Child Psychology*, *119*, 120–126.

Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology*, *131*, 94–103.

Carruthers, P. (2017). Mindreading in adults: evaluating two-systems views. *Synthese*, *194*(3), 673–688.

Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 367-379.

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology*, *33*(5), 781.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*(4), 377–395.

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, *1*(2), 255–259.

Dennett, D. C. (1978). Beliefs about beliefs [P&W, SR&B]. *Behavioral and Brain Sciences*, *1*(4), 568–570.

de Villiers, J. G., & de Villiers, P. A. (2000). *Linguistic determinism and the understanding of false. Children's Reasoning and the Mind* (Eds.), P. Mitchell and K. Riggs (Hove: Psychology Press), 191–228.

de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development*, *17*(1), 1037–1060.

Devine, R. T., & Hughes, C. (2014). Relations between false-belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, *85*(5), 1777–1794.

Drayton, L. A., & Santos, L. R. (2018). What do monkeys know about others' knowledge?. *Cognition*, *170*, 201–208.

Farrar, M. J., Benigno, J. P., Tompkins, V., & Gage, N. A. (2017). Are there different pathways to explicit false-belief understanding? General language and complementation in typical and atypical children. *Cognitive Development*, *43*, 49–66.

Flavell, J. H. (1992). *Perspectives on perspective-taking. Piaget's theory: Prospects and possibilities*. (pp. 107–139). Hillsdale, NJ: Erlbaum

Flavell, J. H. (1993). The development of children's understanding of false-belief and the appearance-reality distinction. *International Journal of Psychology*, *28*(5), 595–604.

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*(2), 109–128.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, *7*(2), 77–83.

Gallagher, S., & Povinelli, D. J. (2012). Enactive and behavioral abstraction accounts of social understanding in chimpanzees, infants, and adults. *Review of Philosophy and Psychology*, *3*(1), 145–169.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193.

Grosse Wiesmann, C., Friederici, A. D., Disla, D., Steinbeis, N., & Singer, T. (2018). Longitudinal evidence for 4-year-olds' but not 2-and 3-year-olds' false-belief-related action anticipation. *Cognitive Development*, *46*, 58–68.

Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. (2017). Implicit and explicit false-belief development in preschool children. *Developmental Science*, *20*(5).

Grosse Wiesmann, C., Schreiber, J., Singer, T., Steinbeis, N., & Friederici, A. D. (2017). White matter maturation is associated with the emergence of theory of mind in early childhood. *Nature Communications*, *8*, 14692.

Gut, A., Haman, M. & Gorbaniuk, O. (under review) The development of understanding opacity in preschoolers: A transition from a coarse- to fine-grained understanding of beliefs.

Gweon, H., Dodell-Feder, D., Bedny, M., & Saxe, R. (2012). Theory of mind performance in children correlates with functional specialization of a brain region for thinking about thoughts. *Child development*, *83*(6), 1853–1868.

Helming, K. A., Strickland, B., & Jacob, P. (2014). Making sense of early false-belief understanding. *Trends in Cognitive Sciences*, *18*(4), 167–170.

Helming, K. A., Strickland, B., & Jacob, P. (2016). Solving the puzzle about early belief-ascription. *Mind & Language*, *31*(4), 438–469.

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging*, *28*(3), 826.

Heyes, C. (2014a). False-belief in infancy: a fresh look. *Developmental Science*, *17*(5), 647–659.

Heyes, C. (2014b). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, *9*(2), 131–143.

Hoehl, S., Wiese, L., & Striano, T. (2008). Young infants' neural processing of objects is affected by eye gaze direction and emotional expression. *PLoS One*, *3*(6), e2389.

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131–134.

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*(4), 590–619.

Hyde, D. C., Aparicio Betancourt, M., & Simon, C. E. (2015). Human temporal-parietal junction spontaneously tracks others' beliefs: A functional near-infrared spectroscopy study. *Human Brain Mapping*, *36*(12), 4831–4846.

Hyde, D. C., Simon, C. E., Ting, F., & Nikolaeva, J. I. (2018). Functional Organization of the Temporal–Parietal Junction for Theory of Mind in Preverbal Infants: A Near-Infrared Spectroscopy Study. *Journal of Neuroscience*, *38*(18), 4264–4274.

Jacoby, N., Bruneau, E., Koster-Hale, J., & Saxe, R. (2016). Localizing pain matrix and theory of mind networks with both verbal and non-verbal stimuli. *NeuroImage*, *126*, 39–48.

Jenkins, J. M., & Astington, J. W. (2000). Theory of mind and social behavior: Causal models tested in a longitudinal study. *Merrill-Palmer Quarterly (1982-)*, 203–220.

Kampis, D., Fogd, D., & Kovács, Á. M. (2017). Nonverbal components of Theory of Mind in typical and atypical development. *Infant Behavior and Development*, *48*, 54–62.

Kovács, Á. M., Kühn, S., Gergely, G., Csibra, G., & Brass, M. (2014). Are all beliefs equal? Implicit belief attributions recruiting core brain regions of theory of mind. *PloS one*, *9*(9), e106558.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*(6012), 1830–1834.

El Kaddouri, R., Bardi, L., De Bremaeker, D., Brass, M., & Wiersema, J. R. (2019). Measuring spontaneous mentalizing with a ball detection task: putting the attention-check hypothesis by Phillips and colleagues (2015) to the test. *Psychological Research*, 1–9.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false-beliefs. *Science*, *354*(6308), 110–114.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2017). A test of the submentalizing hypothesis: Apes' performance in a false-belief task inanimate control. *Communicative & Integrative Biology*, *10*(4), e1343771.

Kuhn, D. (2000). Theory of mind, metacognition, and reasoning: A life-span perspective. In P. Mitchell & K. J. Riggs (Eds.) *Children's reasoning and the mind* (pp. 301–326). Hove: Psychology Press.

Kulke, L., & Rakoczy, H. (2018). Implicit theory of mind – An overview of current replications and non-replications. *Data in Brief*, *16*, 101–104.

Kulke, L., Reiß, M., Krist, H., & Rakoczy, H. (2018). How robust are anticipatory looking measures of theory of mind? Replication attempts across the life span. *Cognitive Development*, *46*, 97–111.

Kulke, L., von Duhn, B., Schneider, D., & Rakoczy, H. (2018). Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychological Science*, 0956797617747090.

Lackner, C., Sabbagh, M. A., Hallinan, E., Liu, X., & Holden, J. J. (2012). Dopamine receptor D4 gene variation predicts preschoolers' developing theory of mind. *Developmental Science*, *15*(2), 272–280.

Lackner, C. L., Bowman, L. C., & Sabbagh, M. A. (2010). Dopaminergic functioning and preschoolers' theory of mind. *Neuropsychologia*, *48*(6), 1767–1774.

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: developingwhat'andwhere'systems. *Trends in Cognitive Sciences*, *2*(1), 10–18.

Lillard, A. (1998). Ethnopsychologies: cultural variations in theories of mind. *Psychological Bulletin*, *123*(1), 3.

Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: relations with complex syntactical mastery. *Child Development*, *81*(2), 597–615

Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive architecture of reasoning in children and adults: A primer on the two-systems account. *Child Development Perspectives*, *10*(3), 184-189.

Mars, R. B., Neubert, F. X., Noonan, M. P., Sallet, J., Toni, I., & Rushworth, M. F. (2012). On the relationship between the "default mode network" and the "social brain". *Frontiers in Human Neuroscience*, *6*, 189.

Meert, G., Wang, J., & Samson, D. (2017). Efficient belief tracking in adults: The role of task instruction, low-level associative processes and dispositional social functioning. *Cognition*, *168*, 91–98.

Meltzoff, A. N. (1993). The role of imitation in understanding persons and developing a theory of mind. *Understanding other minds: Perspectives from Autism*, 335–366.

Moran, J. M., Young, L. L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., & Gabrieli, J. D. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences*, *108*(7), 2688–2692.

Moses, L. J. (2001). Executive accounts of theory-of-mind development. *Child Development*, *72*(3), 688–690.

Naughtin, C. K., Horne, K., Schneider, D., Venini, D., York, A., & Dux, P. E. (2017). Do implicit and explicit belief processing share neural substrates? *Human Brain Mapping*, *38*(9), 4760–4772.

Nelson, K. (1998). *Language in cognitive development: The emergence of the mediated mind*. Cambridge: Cambridge University Press.

Nelson, K., Plesa, D., & Henseler, S. (1998). Children's theory of mind: An experiential interpretation. *Human Development*, *41*(1), 7–29.

Nijhof, A. D., Brass, M., Bardi, L., & Wiersema, J. R. (2016). Measuring mentalizing ability: A within-subject comparison between an explicit and implicit version of a ball detection task. *PloS one*, *11*(10): e0164373. doi:10.1371/journal.pone.0164373

Nijhof, A. D., Bardi, L., Brass, M., & Wiersema, J. R. (2018). Brain activity for spontaneous and explicit mentalizing in adults with autism spectrum disorder: An fMRI study. *NeuroImage: Clinical*, *18*, 475–484.

Oktay-Gür, N., Schulz, A., & Rakoczy, H. (2018). Children exhibit different performance patterns in explicit and implicit theory of mind tasks. *Cognition*, *173*, 60–74.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false-beliefs? *Science*, *308*(5719), 255–258.

Peterson, C., Slaughter, V., Moore, C., & Wellman, H. M. (2016). Peer social skills and theory of mind in children with autism, deafness, or typical development. *Developmental Psychology*, *52*(1), 46.

Peterson, C. C., Slaughter, V. P., & Paynter, J. (2007). Social maturity and theory of mind in typically developing children and those on the autism spectrum. *Journal of Child Psychology and Psychiatry*, *48*(12), 1243–1250.

Phillips, J., Ong, D. C., Surtees, A. D., Xin, Y., Williams, S., Saxe, R., & Frank, M. C. (2015). A second look at automatic theory of mind: Reconsidering Kovács, Téglás, and Endress (2010). *Psychological Science*, *26*(9), 1353–1367.

Poulin-Dubois, D., & Yott, J. (2018). Probing the depth of infants' theory of mind: Disunity in performance across paradigms. *Developmental Science*, *21*(4), e12600.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, *36*(1), 1–16.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Priewasser, B., Rafetseder, E., Gargitter, C., & Perner, J. (2017). Helping as an early indicator of a theory of mind: Mentalism or teleology?. *Cognitive Development*, *46*, 69–78. doi:10.1016/j.cogdev.2017.08.002

Rabinowitz, N. C., Perbet, F., Song, H. F., Zhang, C., Eslami, S. M., & Botvinick, M. (2018). *Machine theory of mind*. arXiv preprint arXiv:1802.07740.

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., & Saxe, R. (2018). Development of the social brain from age three to twelve years. *Nature Communications*, *9*(1), 1027.

Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, *24*(1), 27–33.

Rubio-Fernández, P., Jara-Ettinger, J., & Gibson, E. (2017). Can processing demands explain toddlers' performance in false-belief tasks?. *Proceedings of the National Academy of Sciences*, *114*(19), E3750–E3750.

Sabbagh, M. A., Bowman, L. C., Evraire, L. E., & Ito, J. (2009). Neurodevelopmental correlates of theory of mind in preschool children. *Child Development*, *80*(4), 1147–1162.

Samson, D., Apperly, I., Braithwaite, J., Andrews, B., & Bodley Scott, S. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1255–1266.

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*(2), 65–72.

Scheeren, A. M., de Rosnay, M., Koot, H. M., & Begeer, S. (2013). Rethinking theory of mind in high-functioning autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *54*(6), 628–635.

Schick, B., de Villiers, P., de Villiers, J., & Hoffmeister, R. (2007). Language and theory of mind: A study of deaf children. *Child Development*, *78*(2), 376–396.

Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012). Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science*, *23*(8), 842–847.

Schneider, D., Slaughter, V. P., Becker, S. I., & Dux, P. E. (2014). Implicit false-belief processing in the human brain. *NeuroImage*, *101*, 268–275.

Schneider, D., Nott, Z. E., & Dux, P. E. (2014). Task instructions and implicit theory of mind. *Cognition*, *133*, 43–47.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*, *42*, 9–34.

Schurz, M., Tholen, M. G., Perner, J., Mars, R. B., & Sallet, J. (2017). Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: A review using probabilistic atlases from different imaging modalities. *Human Brain Mapping*, *38*(9), 4788–4805.

Schuwerk, T., Vuori, M., & Sodian, B. (2015). Implicit and explicit theory of mind reasoning in autism spectrum disorders: The impact of experience. *Autism, 19*(4), 459–468.

Scott, R. M. (2014). Post hoc versus predictive accounts of children's theory of mind: A reply to Ruffman. *Developmental Review*, *34*(3), 300–304.

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*, *80*(4), 1172–1196.

Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences*, 21(4), 237–249.

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, *47*(5), 1239.

Slaughter, V. (2015). Theory of mind in infants and young children: a review. *Australian Psychologist*, *50*(3), 169–172.

Sodian, B. (2016). Is false-belief understanding continuous from infancy to preschool age? In D. Barner & A. S. Baron (Eds.), *Core knowledge and conceptual change* (p. 301), Oxford: Oxford University Press.

Sommerville, J. A., Bernstein, D. M., & Meltzoff, A. N. (2013). Measuring beliefs in centimeters: Private knowledge biases preschoolers' and adults' representation of others' beliefs. *Child Development*, *84*(6), 1846–1854.

Southgate, V. (2018). The puzzle of early mentalizing: A unitary theory of infants' successes and preschoolers' failures. doi:10.17605/OSF.IO/Y2GHQ

Spreng, R. N., & Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *Journal of Cognitive Neuroscience*, *22*(6), 1112–1123.

Stojnić, G. (2017). *Can two-year olds understand others' false-beliefs about object identities?* (Doctoral dissertation, Rutgers University-Graduate School-New Brunswick).

Stojnić, G. & Leslie, A. M. (2018). Poster presented at 9th Budapest Conference on Cognitive Development, Budapest, January 2018.

Thoermer, C., Sodian, B., Vuori, M., Perst, H., & Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false-belief from infancy to preschool age. *British Journal of Developmental Psychology*, *30*(1), 172–187.

Tomasello, M. (1995). Joint attention as social cognition. *Joint Attention: Its Origins and Role in Development*, 103–130.

Tomasello, M. (2018). How children come to understand false-beliefs: A shared intentionality account. *Proceedings of the National Academy of Sciences*, *115*(34), 8491–8498.

Wang, L., & Leslie, A. M. (2016). Is Implicit Theory of Mind the 'Real Deal'? The Own-Belief/True-Belief Default in Adults and Young Preschoolers. *Mind & Language*, *31*(2), 147–176.

Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford: Oxford University Press.

Wellman, H. M. (2017). The development of theory of mind: Historical reflections. *Child Development Perspectives*, *11*(3), 207–214.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523–541.

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, *35*, 245–275.

Wellman, H. M., Fang, F., & Peterson, C. C. (2011). Sequential progressions in a theory-of-mind scale: Longitudinal perspectives. *Child Development*, *82*(3), 780–792.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Yott, J., & Poulin-Dubois, D. (2016). Are infants' theory-of-mind abilities well integrated? Implicit understanding of intentions, desires, and beliefs. *Journal of Cognition and Development*, *17*(5), 683–698.

Zaitchik, D. (1990). When representations conflict with reality: The preschooler's problem with false-beliefs and "false" photographs. *Cognition*, *35*(1), 41–68.

Zelazo, P. D., Carlson, S. M., & Kesek, A. (2008). The development of executive function in childhood. In C. A. Nelson & M. Luciana (Eds.), *Developmental cognitive neuroscience. Handbook of developmental cognitive neuroscience* (pp. 553–574). Cambridge, MA, US: MIT Pres