Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience
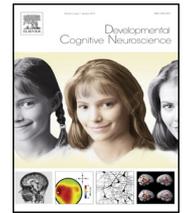
Review

How and where: Theory-of-mind in the brain☆

Caitlin E.V. Mahy [a,b,∗], Louis J. Moses [a,1], Jennifer H. Pfeifer [a,1]

[a] University of Oregon, USA
[b] University of Geneva, Switzerland

ARTICLE INFO

ABSTRACT

Theory of mind (ToM) is a core topic in both social neuroscience and developmental psychology, yet theory and data from each field have only minimally constrained thinking in the other. The two fields might be fruitfully integrated, however, if social neuroscientists sought evidence directly relevant to current accounts of ToM development: modularity, simulation, executive, and theory theory accounts. Here we extend the distinct predictions made by each theory to the neural level, describe neuroimaging evidence that in principle would be relevant to testing each account, and discuss such evidence where it exists. We propose that it would be mutually beneficial for both fields if ToM neuroimaging studies focused more on integrating developmental accounts of ToM acquisition with neuroimaging approaches, and suggest ways this might be achieved.

© 2014 The Authors. Published by Elsevier Ltd. All rights reserved.

Contents

∗ Corresponding author at: Department of Psychology, 1227 University of Oregon, Eugene, OR 97403, USA. Tel.: +1 541 346 4881.
E-mail address: cmahy@uoregon.edu (C.E.V. Mahy).
[1] Department of Psychology, 1227 University of Oregon, Eugene, OR 97403, USA.

1. Introduction

As the field of psychology diversifies, communication among researchers who study the same topic with different methodologies becomes increasingly important. Empirical findings are disseminated within a certain theoretical or methodological framework, potentially creating gaps between literatures that may not be bridged. Such is the current state of theory of mind (ToM) research, where there is little overlap in how developmental psychologists and social neuroscientists study the ways in which people impute mental states to self and other.

Social neuroscience and developmental psychology both prominently feature research on ToM, yet emphasize different facets of this core social cognitive ability. Social neuroscientists tend to focus on *where* in the brain mentalizing resides, while developmental psychologists are centrally concerned with *how* mentalizing is acquired (and *when* it emerges). As a result, much of the adult neuroimaging work has not been explicitly linked to developmental theories; instead, it has focused on identifying ToM-relevant neural regions, and distinguishing groups of individuals such as those with and without autism spectrum disorders in terms of systems or processes recruited. The differing research agendas of these two fields are clearly related and yet, as others have noted (e.g., Apperly, 2008; Saxe, 2006), they remain largely unreconciled.

In this paper we aim to advance the field by attempting to integrate developmental and neuroimaging approaches to ToM. We propose that it would be mutually beneficial for both fields if developmental and social neuroscientists were to more fully consider developmental theories focusing on acquisition of ToM when generating hypotheses, designing studies, and interpreting results. Our review is the first to examine in one place neural evidence for the four major theories of ToM acquisition that have been proposed, although others have compared subsets of the theories (e.g., Apperly, 2008; Wilkinson and Ball, 2012). We take the position that processes specified by each theory may potentially contribute to ToM development and that ultimately neuroimaging research may help generate a new theory that integrates existing approaches. Our primary goal is to compile an up-to-date summary of neuroimaging evidence relevant to theoretical accounts of ToM acquisition so that this growing field may advance conceptually, theoretically, and methodologically. Further, we hope to establish where neuroimaging techniques might be particularly helpful or unhelpful in testing a given theory. It is our hope that both behavioral and neuroimaging researchers will find this review useful and that it will stimulate future work integrating developmental and neuroimaging approaches.

We begin by briefly reviewing the main theoretical accounts of how ToM is acquired and suggesting the types of neural evidence that would support or pose problems for each theory, describing such evidence where it is available. We consider the following four development accounts of ToM: modularity theories, simulation theories, executive accounts, and theory theory. We draw on developmental evidence whenever possible although the current state of the literature dictates a heavy reliance on findings from adult studies, as neuroimaging studies with children remain scarce. We also draw on the autism literature to the extent it is strongly germane to our argument, although a full consideration of that literature is beyond the scope of the current paper (see Happé and Frith, 2013 for a review of neuroimaging studies in ASD). Finally, we address some of the challenges of using neuroimaging techniques to examine developmental theories of ToM and suggest future research that could close the gap between social neuroscience and developmental approaches.

2. Modularity theories: selectivity

2.1. Definition

Modularity theories (e.g., Fodor, 1983; Leslie et al., 2004; Scholl and Leslie, 1999; Baron-Cohen, 1995, 1998) postulate that ToM development is driven by an innate neural mechanism dedicated to mental state reasoning. Although experience may be important in triggering this mechanism, it cannot revise the mechanism's basic nature. Leslie and his colleagues have proposed the most fully articulated and prominent modularity theory of ToM and therefore we focus on their account. It should be noted, however, that their account represents a relatively strong form of modularity and that accounts stipulating less stringent criteria have been proposed (e.g., see Coltheart, 1999; Carruthers, 2003). The central claims made by Leslie and colleagues are that an innate ToM module (ToMM) is working by the second year of life, and that later age-related improvements in ToM performance in childhood are driven by an inhibitory selection process that becomes increasingly able to handle the executive demands of ToM tasks (Leslie et al., 2004; Scholl and Leslie, 2001; German and Hehman, 2006).

2.2. Criteria

An early developing ToMM is consistent with findings from studies suggesting that even infants may be capable of at least some mental state inferences (e.g., Baillargeon et al., 2010; Kovacs et al., 2010; Onishi and Baillargeon, 2005), while the proposed selection process may account for performance improvements in explicit, verbal ToM tasks commonly used with preschoolers (e.g., false belief tasks; Leslie and Polizzi, 1998). Convincingly demonstrating modularity, however, is anything but straightforward. Domain specificity, obligatory firing, rapid speed of processing, constrained output, dedicated neural architecture, and a characteristic pattern of breakdown are typically associated with modular processing (Baron-Cohen, 1994; Coltheart, 1999; Fodor, 1983; Baldwin and Moses, 1994; Scholl and Leslie, 1999). Given these stringent criteria, Leslie et al.'s modularity theory is perhaps the most falsifiable of all the ToM accounts.

Neuroimaging evidence is particularly useful in testing this theory as clear predictions can be generated from the purported existence and developmental maturation of a ToMM. First, to support modularity, a particular brain region, or network of brain regions, would need to be consistently activated whenever individuals engage in mental state reasoning. Second, this pattern of activation would need to be present not only early in development but throughout the lifespan. Recruitment of the neural network could be modulated by experience, but the same set of neural regions should be consistently engaged at all points in development. For example, evidence of early ToM competence in infancy would more strongly support a modularity account, if infants in fact recruit similar neural regions as adults when mentalizing. Third, to satisfy the dedicated neural architecture criterion, the neural network would also need to be selectively recruited for ToM reasoning. If the neural system responsible for mental state reasoning is not selective, or only becomes selective to mental states late in childhood or in adulthood, that would challenge the modularity account. Finally, if Leslie et al.'s (2004) conception of an inhibitory process is correct, developmental improvements in ToM during the preschool period should be associated with a neural region related to an inhibitory selection process rather than with a neural region related to ToM.

2.3. Evidence

Consistent with modularity accounts, a common neural network involved in adults' mental state reasoning has been established, although the components of the network vary to some extent from study to study. Brain regions most typically implicated include (a) cortical midline structures (CMS) comprised of the medial prefrontal cortex (MPFC), adjacent rostral anterior cingulate cortex (rACC), and medial posterior parietal cortices (MPPC) including posterior cingulate and precuneus (Amodio and Frith, 2006; Mitchell et al., 2005; Ochsner et al., 2005), and (b) the bilateral temporal parietal junction (TPJ; Saxe, 2009; Young et al., 2010a).

Because the CMS also support abilities such as integration and prospection (e.g., Spreng et al., 2009; Spreng and Grady, 2010), and hence do not respond selectively to ToM input, these structures do not qualify as a ToMM. However, the TPJ appears to be a stronger candidate for a ToMM as some evidence suggests that, unlike the CMS, it may selectively respond to mental state information (Saxe et al., 2009; Samson et al., 2004).

Early neuroimaging studies found that bilateral TPJ was recruited during mental state reasoning in adults (Saxe and Kanwisher, 2003; Saxe and Powell, 2006). Appropriate comparison conditions with respect to selectivity are key to the TPJ as ToMM argument. For example, Saxe and Kanwisher (2003) found that the bilateral TPJ was recruited more heavily when participants listened to stories about characters' mental states compared to stories about characters' physical descriptions in otherwise identical formats (see also Saxe et al., 2009). Further, Saxe and Powell (2006) found that bilateral TPJ and posterior cingulate were recruited during stories about a protagonist's thoughts but not during stories about a protagonist's non-mental internal states such as bodily sensations or physical attributes such as appearance.

In recent years the right TPJ has received particular scrutiny as it appears to be more responsive to thinking about mental states than the left TPJ (Dohnel et al., 2012; Saxe, 2010), although some have argued the left TPJ is also necessary for representing others' beliefs (e.g., Samson et al., 2004). Some evidence in favor of a special role for the right TPJ comes from an EEG/ERP investigation with adults, in which belief reasoning recruited right posterior neural systems (Liu et al., 2009). In addition, specifically disrupting the functioning of right TPJ using transcranial magnetic stimulation (TMS) has been found to increase participants' difficulty in using mental state information in moral judgments (Young et al., 2010b). More recently, Dohnel et al. (2012) also found that the right TPJ was activated in both true and false belief reasoning in adults. Taken together, studies with adults such as these support the argument that the TPJ, and the right TPJ in particular, may be selective for mental state reasoning and may therefore be a plausible candidate for a ToMM.

As noted earlier, however, TPJ should also demonstrate similarly selective properties in children and adolescents in order to qualify as a ToMM. Here the evidence is not as strongly supportive of modularity. Although 8–12-year-old children significantly engage TPJ bilaterally in ToM stories that require second-order false belief reasoning compared to a non-ToM control story (Kobayashi et al., 2007), younger children do not always show this selective TPJ recruitment. For instance, in a sample of 6–11-year-olds, the younger children showed a lack of selective TPJ recruitment during ToM reasoning compared to more general social reasoning (Saxe et al., 2009). These children recruited right TPJ equally for mental and physical facts about people, whereas the older children engaged right TPJ only for mental facts. Relatedly, Gweon et al. (2013) found increasing selectivity to mental state information with age in TPJ bilaterally in a sample of 5–11-year-olds. Overall, this region appears to become increasingly selective for mental state information during development. Such a developmental trajectory

poses a problem for current modularity theories that are based on the assumption that a selective ToMM is in place much earlier in life.

Moreover, even in adults, tasks unrelated to ToM reasoning also appear to recruit TPJ, again violating the specificity criterion. Sabbagh (2011) has suggested that right TPJ may be 'shared real estate' supporting *processes* that underlie both ToM and other tasks. For example, the right TPJ is recruited during attentional reorienting (Mitchell, 2008), focused attention and target detection (Geng and Mangun, 2011; Geng and Vossel, 2013; Kubit and Jack, 2014), and in motor response inhibition (Rothmayr et al., 2010).

It remains possible, however, that more specific sub-regions of the right TPJ are dedicated to ToM as distinct from the other processes just mentioned. For example, Scholz et al. (2009) found that ToM and exogenous attention activated different areas of the right TPJ. Further, Rothmayr et al. (2010) found that dorsal aspects of the left TPJ might be involved in attentional reorienting whereas more ventral aspects may be dedicated to belief-attribution. The heterogeneity of these results surrounding the role of right versus left TPJ and lack of specificity within the TPJ in ToM processes may be due to fMRI data acquired at different resolutions, as higher resolutions will reveal more specific or distinct activations. At any rate, it remains unclear at this point whether the TPJ or parts of the TPJ are selective in reasoning about mental states or whether there is a common process underlying ToM and attentional orienting that TPJ supports. When combined with the evidence described earlier concerning inconsistent engagement of the TPJ in young children's ToM reasoning (Saxe et al., 2009), the absence of clear evidence of specificity for mental state reasoning comes close to ruling out the view that TPJ constitutes a theory of mind module. Of course, modularity theories that do not propose TPJ as the 'ToMM' would not be challenged in these ways. At this point, however, a compelling argument for an alternative region as the neural basis for ToMM is yet to be made.

3. Simulation theories: shared neural representations

3.1. Definition

Simulation theories (Gallese and Goldman, 1998; Goldman, 2009; Gordon, 1992; Harris, 1992, 2000) propose that children (and adults) rely upon direct access to their own psychological states in order to make mental state attributions. In reasoning about the minds of others, they project into others' "shoes" and then read off what they would experience in the relevant situation. Simulation theories nicely account for the extended developmental trajectory of ToM reasoning from the early understanding of false beliefs to later success on more challenging ToM tasks that require more complex simulation (Carpendale and Lewis, 2004; Schwanenflugel et al., 1996). They are also consistent with work demonstrating relations between children's ToM and imaginative ability, which is argued

to underpin simulation (e.g., Lillard and Kavanaugh, 2014; Taylor and Carlson, 1997).

3.2. Criteria

Evidence in favor of simulation accounts would need to show that neural systems underlying imaginative or mental simulation processes are involved in children's ToM reasoning as well as in mapping of the self to others. In addition, children's experience necessarily plays a more central and formative role in this theory than in modularity accounts as it is practice in perspective-taking that is argued to improve simulation skills and thus ToM (Chandler, 1973; Harris, 1992; Ozonoff and Miller, 1995). Hence, over development, these neural systems should become more efficient and automatized. Theoretically, simulation processes should be able to be studied with neuroimaging techniques to the extent that they rely on particular neural networks whose strength of activation is related to ToM reasoning based on the similarity of another person to the self.

3.3. Evidence

3.3.1. Cortical midline structures

Two candidate neural systems that may provide two routes to simulation through the mapping of self to other have been discussed. The first system includes the cortical midline structures (CMS). CMS may support evaluative simulation (Uddin et al., 2007) as these structures have been suggested to be involved in intentional, controlled, evaluative thoughts about self and other. Recruitment of these structures could be indicative of simulation processes because they are involved both in self-perception and perspective taking, at least in adolescents and adults (Pfeifer et al., 2009; Saxe et al., 2006; Spreng et al., 2009). For example, Saxe et al. (2006), found activation of CMS such as MPFC and medial precuneus in both a self-reflection task and a belief reasoning task in adults.

It is important to note that studies investigating the CMS have often focused not on others' mental states, but rather on appraisals of the self and others (see Pfeifer et al., 2009). For example, rather than directly asking participants to think about another's thoughts (e.g., "where does she think the ball is?"), the studies in question tend to ask individuals what they themselves think about their own or another's traits (e.g., "are you smart?" or "is your friend smart?"). The relevance of this research, however, is that individuals seem to activate similar brain areas when evaluating others and the self.

Behavioral work shows individuals are more likely to project their mental states onto others who are perceived as similar than onto those who are perceived as dissimilar (Ames, 2004a,b). Consistent with those findings, two sub-regions of the MPFC, specifically its ventral (lower) and dorsal (upper) aspects, are differentially active in adults when thinking about a similar or dissimilar other: the ventral MPFC tends to be more responsive during mentalizing when the other person is more similar to the self whereas the dorsal MPFC is more responsive when making social judgments about dissimilar others (Mitchell et al.,

2005, 2006). Further, Saxe and Wexler (2005) found a trend toward higher activity in dorsal than ventral MPFC when participants considered the mental states of individuals with a foreign compared to a familiar background. This body of evidence is supported by a recent meta-analysis that revealed a spatial gradient for mentalizing in the MPFC (Denny et al., 2012) such that self-related judgments were associated with activity in relatively ventral MPFC, whereas other-related judgments were associated with activity in relatively dorsal MPFC. In sum, sub-regions of the MPFC differentiate between thinking about the attributes and/or mental states of similar versus dissimilar others, suggesting that certain regions may be playing a role both in thinking about the self and in simulating mental states for similar others, whereas other regions may support a different kind of processing in order to make judgments about a dissimilar other (Harris, 1992; Nichols et al., 1995).

The findings discussed thus far suggest a potential role for the CMS in online simulation, particularly in adults. Beyond that, what evidence is there that the CMS plays a role in the *acquisition* of ToM during childhood? Although adult studies have shown the dorsal/ventral specialization in MPFC when evaluating similar and dissimilar others, studies examining this distinction in children are scarce. However, in a study comparing 9- and 10-year-olds' and adults' evaluations of the self and a familiar, fictional other (Harry Potter), activity in the MPFC changed with development (Pfeifer et al., 2007). Specifically, as in the studies just discussed, adults recruited dorsal MPFC for reasoning about Harry Potter (arguably a fairly dissimilar other), but ventral MPFC for reasoning about themselves. Like the adults, children recruited dorsal MPFC for reasoning about Harry Potter but, in contrast to adults, they recruited both dorsal and ventral MPFC for reasoning about themselves. This change in recruitment of MPFC may suggest that adults have specialized processes for simulating and reasoning about dissimilar (and similar) others, whereas children's brains may have not yet specialized in this manner.

3.3.2. Mirror neuron system

The second candidate neural substrate for simulation processes is the putative mirror neuron system (MNS), which co-activates to the actions, intentions, and emotions of both the self and others in adults and children (Dapretto et al., 2006; Iacoboni, 2009; Iacoboni et al., 2005; Iacoboni and Dapretto, 2006; Pfeifer et al., 2008; Rizzolatti and Craighero, 2004; Waytz and Mitchell, 2011). The MNS is a network composed of regions in the inferior frontal gyrus (pars operculerus and adjacent ventral premotor cortex) and rostral part of the inferior parietal lobule that are recruited during the perception and execution of identical actions (Dapretto et al., 2006; Iacoboni and Dapretto, 2006). This network may aid in simulating others' mental states by enabling a direct mapping of others' actions, goals, and intentions to the self. MNS activity is present in both non-human and human primates (Gallese, 2007). Because much of the work with non-human primates has taken advantage of single cell recordings, the MNS is better characterized at a neural level in those species.

Given that we share this system with our recent primate ancestors, and that it has been documented in adults and older children (see Iacoboni and Dapretto, 2006 for a review), it is worth examining evidence for an early developing MNS in human infants. The presence of such a system would suggest that simulation processes are operating early in life. Behavioral research suggests that 10–11-month-old infants can parse the structure of human action, a prerequisite for understanding intentional action (Baldwin et al., 2001; Saylor et al., 2007), and that infants under one year of age appreciate the intentionality in others' actions (e.g., Behne et al., 2005; Meltzoff, 2007; Phillips et al., 2002). A mirror system that maps others' intentional actions with reference to one's own actions could play an important role in the acquisition of concepts such as intention, desire, and belief via simulation. As with modularity accounts, it would need to be established that the mirror neuron system (or the CMS) is involved in very early ToM reasoning in infants and young children before concluding that simulation is central to the *acquisition* of ToM (as opposed to online reasoning in a more mature ToM).

Studies using EEG technology may help to identify MNS function in infants. Mu rhythm, an electrical signal around 10 Hz generated by motor areas at rest, is suppressed both during motor execution and action observation in adults (Glenberg, 2011). Infant EEG mu rhythm is a possible candidate for an early form of the mirror neuron system, as desynchronization of mu rhythm occurs during an infant's own movement as well as the observation of another's movement (Marshall et al., 2010; Southgate et al., 2009, 2010). Research on infant mu rhythm is still in its early stages and so we cannot be fully confident that mu rhythm desynchronization represents a mirroring process related to the MNS. It will be important to determine whether mu rhythm response changes with development, and whether it performs the same function as the MNS, before making the claim that it supports simulation theory (Marshall and Meltzoff, 2011).

In sum, there are at least two neural systems that could potentially support ToM reasoning through the simulation of others' mental states. Additional support for the role of CMS and the MNS in ToM reasoning comes from research indicating that individuals with autism show reduced activation of these networks in social cognition tasks compared to typically developing controls (Dapretto et al., 2006). Because these individuals have clear deficits in ToM, the findings are consistent with the view that these two systems play a role in successful mentalizing (Dapretto et al., 2006; Pfeifer et al., 2011, 2013; Uddin et al., 2007). It is worth noting, however, that abnormalities in MNS and CMS functioning in individuals with autism may be related to other difficulties that this population experiences, such as linguistic or cognitive deficits, and may not be the sole reason for their difficulty in understanding the minds of others via simulation (or any other process for that matter).

4. Executive accounts: role of inhibition

4.1. Definition

Executive accounts (e.g., Carlson et al., 2014; Carlson and Moses, 2001; Hughes, 1998; Moses, 2001; Moses and Tahiroglu, 2010; Russell, 1997) posit that children's

difficulties reasoning about mental states stem at least in part from challenges in inhibiting one's own perspective in order to generate a different one, and/or in holding the relevant perspectives in working memory. These accounts postulate that developments in executive functioning (EF) contribute markedly to age related changes in ToM during the preschool years, and perhaps beyond. At least two possible stances might be taken regarding the role of EF in ToM reasoning. One strong stance might be that EF is *sufficient* for ToM, meaning that EF can support ToM reasoning in the absence of additional processes. A second, weaker stance might be that EF is not sufficient for ToM (that is, it relies on other processes as well) but that it is nonetheless *necessary* for ToM. This weaker stance itself takes two forms according to which EF is necessary for either (a) the expression of ToM concepts in online task performance or (b) the emergence or acquisition of ToM concepts themselves (Moses, 2001).

4.2. Criteria

Executive accounts are supported by a plethora of findings demonstrating strong relations between EF and ToM that hold up when age, verbal ability, and other possible confounding variables are controlled (see Devine and Hughes, in press, for a recent meta-analysis). To the extent that EF contributes to ToM reasoning and its development, the brain regions that support EF should be engaged during ToM tasks. Further, evidence that brain regions involved in ToM and EF are proximal or structurally connected, and functionally connected during ToM reasoning, would be supportive of such accounts. In principle, neuroimaging evidence should be informative as connectivity between neural regions supporting EF and ToM can be easily examined using MRI.

To support the stance that EF is sufficient for mentalizing, ToM tasks should activate EF areas but not additional ToM-specific brain regions. To support the stance that EF is necessary for ToM, ToM tasks should always activate EF areas but could also activate ToM specific regions. In the following section, we review the existing neuroimaging evidence for these two viewpoints.

We also focus on one particular EF, inhibitory control, for several reasons: (1) inhibitory control has been suggested as a key process in ToM development during the preschool years (e.g., Carlson et al., 1998, 2002, 2004; Carlson and Moses, 2001; Russell, 1997), (2) behavioral and neuropsychological research with adults shows that inhibition plays a necessary but not sufficient role in certain ToM computations (see Apperly et al., 2004; Dumontheil et al., 2010; Fanning et al., 2012; Keysar et al., 2003; Samson, 2009; Stone et al., 1998 for critical complementary perspectives from the neuropsychology literature), and (3) the neural correlates of inhibition have been clearly mapped in the brain. For example, activity in the bilateral ventral prefrontal cortex, right parietal lobe, and right dorsolateral prefrontal cortex increases during inhibition tasks in both children and adults (e.g., Aron et al., 2004; Bunge et al., 2002; Casey et al., 1997; Durston et al., 2002; Fassbender et al., 2006; Levy and Wagner, 2011; Yamaguchi et al., 2008).

4.3. Evidence

A first line of relevant evidence comes from studies with adults comparing neural activation on ToM tasks to that on inhibition tasks. For example, Van der Meer et al. (2011) found that both a high inhibition false-belief task and a stop-signal inhibition task recruited bilateral inferior frontal gyrus, possibly reflecting a common inhibitory control mechanism. In addition, Rothmayr et al. (2010) found that inhibitory control and false belief reasoning tasks recruit some similar neural regions (right superior and medial frontal gyrus, right middle temporal gyrus, bilateral middle frontal gyrus, and bilateral TPJ). Challenging the position that EF is sufficient for ToM, however, Rothmayr et al. also found that inhibition and false belief tasks recruited some distinct neural regions (left superior and medial frontal gyrus, left inferior, middle, and superior frontal gyrus, left middle temporal gyrus, left TPJ, bilateral precuneus, and bilateral thalamus).

Nonetheless, studies such as these are not well suited to uncovering either the *unique* contribution of inhibition to mentalizing or the distinct neural signatures of the two abilities because the tasks differ in so many other ways such as verbal demands, number of trials, and task structure. Clearly, more stringent controls are necessary.

A second line of evidence attempts to provide such control by comparing neural underpinnings of tasks assessing understanding of mental representations (such as false beliefs) with those assessing understanding of non-mental representations (such as "false" photographs and "false" signs). For example, Saxe and Kanwisher (2003) found that false belief stories activated bilateral TPJ in adults to a much greater extent than stories that involved false photographs. This was taken as evidence that ToM recruits neural regions over and above those involved in inhibition as both false belief and false photograph conditions appear to pose identical demands on inhibitory control, only differing in their demands on mentalizing (although see Sabbagh et al., 2006). Further supporting a distinction between the neural systems supporting ToM and EF is a study showing that the right TPJ was specifically associated with processing mental states such as false beliefs, whereas the left TPJ was also activated by reasoning about false signs (Perner et al., 2006). Similar to Saxe and Kanwisher's (2003) study, the false belief and false sign conditions were designed to be identical in their demand on inhibition, yet right TPJ additionally supported false belief reasoning, suggesting that EF is not sufficient for mental state reasoning.

A third line of evidence attempts to uncover common and unique neural underpinnings by manipulating the belief task itself. For example, in the Van der Meer et al. (2011) study discussed earlier two false belief conditions were contrasted, one requiring false belief reasoning without prior knowledge of the object's location (low inhibition condition) and one requiring false belief reasoning with prior knowledge of that location (high inhibition condition). Belief reasoning in the high inhibition compared to the low inhibition condition more heavily recruited areas associated with cognitive control or conflict monitoring such as bilateral inferior frontal gyrus, dMPFC, and insula. Similarly, Hartwright et al. (2012) manipulated whether

the belief to be processed was true (low inhibition) or false (high inhibition). Across conditions they found modulated brain activity in the ToM network in adults (e.g., bilateral TPJ) but also in executive control regions (e.g., ventrolateral PFC). These findings provide evidence suggesting that EF is necessary for false belief reasoning but not true belief reasoning and that different neural regions supporting executive control may be recruited for different aspects of ToM reasoning (i.e., for true and false beliefs).

Taken together, current evidence suggests that although EF (or at least inhibitory processes) is involved in many types of ToM reasoning (particularly when the task requires false belief reasoning), there are distinct neural correlates of ToM reasoning that do not involve EF brain areas. Adult neuroimaging evidence thus seems to rule out the 'strong' account of EF as sufficient for ToM. However, evidence for the weaker 'EF as necessary for ToM' viewpoint is supported by most of the findings showing that neural regions supporting inhibition are often involved in false belief reasoning.

Nonetheless, as is true for the theoretical accounts discussed earlier, the evidence is limited in what it can tell us about the role of EF in ToM *acquisition* during childhood. To our knowledge, there is only one neuroimaging study examining the EF and ToM relation in children. Sabbagh et al. (2009) examined relations between ToM and preschoolers' resting alpha (6–9 Hz), an EEG waveform thought to reflect functional, domain-general, maturational changes in brain development (Thatcher, 1992). Resting EEG alpha estimates in dorsomedial PFC and TPJ were positively associated with individual differences in false belief understanding independent of variability in age and performance on EF tasks. The findings suggest that EF is not sufficient for ToM early in development, although it leaves open the question of whether EF is necessary for, or facilitative of, the expression of ToM earlier in development and throughout the lifespan (see Apperly et al., 2009; Samson et al., 2004). Examining whether brain functioning during EF tasks explains unique variance in ToM performance beyond age and verbal intelligence would be a positive next step.

Developmental studies examining the neural basis of EF and ToM in young children will be critical in helping tease apart the expression and emergence accounts of the role of EF in ToM. As described earlier, the expression account suggests that executive processes are necessary for online performance on ToM tasks whereas the emergence account suggests that ToM concepts cannot be formed without a certain level of executive ability (Moses, 2001). If a mature pattern of recruitment of ToM neural regions but not EF regions is apparent, and yet behavioral ToM performance is immature, this would support the expression account. In contrast, if the functional development of brain regions associated with ToM lagged behind development of EF regions, this would be consistent with the emergence account.

Other analytic techniques may continue to help to identify common circuitry underlying EF and ToM. In particular, a conjunction analysis (Nichols et al., 2005) requires that a particular voxel be significantly activated above the selected statistical threshold in all conditions of interest. Enacting this approach would at least allow for the examination of which systems may be involved in both EF and ToM processing, as demonstrated by Rothmayr et al. (2010). Further, exploring the connectivity among regions during different tasks using psychophysiological interaction (PPI) analysis (e.g., Friston et al., 1997) may be helpful in revealing similarities and differences between ToM and EF processing. For example, during mental state tasks (but not executive, non-mentalizing conditions) there should be connections between bilateral inferior frontal gyrus, the cortical midline structures, and TPJ. Parametric modulation (see Durston and Casey, 2006) is yet another technique that may be helpful in disentangling the role of EF in ToM by allowing for the manipulation of executive demands from low to high in a ToM task (also see Van der Meer et al., 2011), which would potentially reveal ToM regions that are responsive to EF demands.

To sum up, some neural evidence supports executive accounts of ToM development. However, while there is overlap between regions that support ToM and EF, it is once again less clear what role these areas play in *acquisition* of ToM concepts. In line with behavioral findings, the neural evidence suggests that EF is not sufficient for ToM but rather that EF may be necessary for some types of ToM reasoning.

5. Theory theory: conceptual change

5.1. Definition

Theory theory (Gopnik, 2003; Gopnik and Wellman, 1994, 2012) postulates that knowledge about the mind resides in domain-specific theory-like structures and that radical conceptual changes drive the development of children's naïve mental state understanding. According to this account, children collect evidence about the relation between mental states and action, much as a scientist collects data to inform theory. To the extent that such evidence is inconsistent with children's current theory of mind, conceptual change will eventually occur. This theoretical stance suggests that relatively abstract theorizing about data gleaned from the social world forms a system of mental concepts; therefore, the child and his or her experiences play an active role in concept formation (e.g., Cutting and Dunn, 1999; Hughes and Leekham, 2004; Jenkins and Astington, 1996; Lillard and Kavanaugh, 2014; Pears and Moses, 2003).

5.2. Criteria

In the developmental psychology literature, it has often been suggested that *theory theory* provides the best account of existing developmental data, notably progressions in children's appreciation of simpler to more complex mental state concepts (Gopnik and Wellman, 1994; Wellman et al., 2001; Moses, 2001) There is also ample behavioral evidence that children make conceptual advances in mental state understanding in response to experience (e.g., Astington and Baird, 2005; Jenkins and Astington, 1996; Lohmann and Tomasello, 2003; Ruffman et al., 2002; Slaughter and Gopnik, 1996).

Theory theory could be supported by evidence of (a) neural structures or networks supporting domain-specific conceptual structures and (b) age-related neural changes associated with relevant conceptual changes. The central prediction of theory theory, that conceptual change in response to experience drives the development of ToM, is challenging to assess, however, because it is not entirely clear how conceptual change is reflected in the brain. As a result, theory theory may be the most challenging account to evaluate with neuroimaging evidence.

5.3. Evidence

With respect to knowledge about the mind residing in domain-specific structures, much of the evidence for modularity theories becomes relevant. For example, evidence suggesting that TPJ specifically supports ToM is consistent with theory theory as well as modularity accounts. Moreover, the increasing developmental selectivity of the TPJ to ToM concepts that is problematic for modularity accounts is not damaging to theory theory, in that increasing selectivity may reflect conceptual change with advancing age. In addition, while lack of selectivity of brain regions (such as the TPJ) is problematic for modularity theories (because modularity theories postulate dedicated neural architecture), it is not a strike against theory theory, which is neutral with respect to whether the same neural system could be recruited for different purposes.

With respect to the role of radical conceptual change in ToM development, the TPJ may be implicated. For example, recent research with adults suggests that the TPJ may play a role in disengaging from one's internal, current self-perspective to attend to an external, other's point of view (Corbetta et al., 2008) – a process that is key to most ToM tasks. Further, disruption of TPJ activity by seizures or electrical stimulation results in hallucinatory misperceptions surrounding one's body and the environment (see Lenggenhager et al., 2007; Arzy et al., 2006), again suggesting that the TPJ plays a role internal versus external representations. If TPJ is indeed involved in updating one's internal expectations versus external reality, it is a reasonable hypothesis that this brain region may play a role in driving conceptual understanding, as children must integrate information from their external environment with their internal states in order to update their hypotheses.

Relevant to this hypothesis are recent developmental studies using a load force adjustment paradigm (Sabbagh, 2011; Sabbagh et al., 2010), a task that requires individuals to adapt their lifting behavior to smoothly lift an unexpectedly heavy object. The load force adjustment task measures how efficiently children adapt their conception of an object's weight in response to empirical evidence that conflicts with an initial hypothesis. Importantly, an earlier fMRI study with adults found that the rTPJ was engaged during a load force adjustment task suggesting that it is involved in the process of updating prior hypotheses about weight toward more accurate predictions (Jenmalm et al., 2006), similar to how one would update hypotheses for ToM according to theory theory. Consistent with theory theory,

Sabbagh et al. found that ToM reasoning was related to 3- and 4-year-old children's ability in load force adjustment, arguably because both ToM and load force adjustment require change to existing conceptual structures in response to experience. Further, children's resting EEG alpha in TPJ was related to both superior ToM reasoning *and* load force adjustment, suggesting that common neural development associated with conceptual change may be driving performance on these two tasks (Sabbagh, 2011).

Within the theory theory tradition, much attention has been paid to the different developmental trajectories of understanding desires (a relatively early developing concept) and beliefs (a relatively later developing concept; Wellman and Liu, 2004). It is possible that belief understanding develops out of and is supported by desire understanding. If so, belief and desire reasoning may share some underlying neural correlates but belief reasoning should recruit additional neural structures. Consistent with this possibility, adult reasoning about desires and beliefs in an ERP study had both common and unique neural underpinnings. Reasoning about both beliefs and desires was associated with a positive late slow wave with a midfrontal scalp distribution, whereas only reasoning about beliefs showed an additional positive late slow wave with a right-posterior scalp distribution (Liu et al., 2009). This pattern was replicated by Bowman et al. (2012) in 7–8-year-old children. Therefore, it seems plausible that early desire reasoning relies on neural regions and perhaps conceptual structures that in part support later belief understanding.

A potential approach to gaining further traction on ToM theory change would be to examine whether ToM relevant brain structures/systems such as the CMS, MNS, and TPJ show developmental changes associated with advancing ToM understanding. Perhaps the system that changes the most over the early childhood years or shows evidence of connectivity changes is the system that plays a role in driving conceptual change. Similarly, changes in white matter connectivity, which are known to be refined by experience (Keller and Just, 2009; Mabbott et al., 2006; Nagy et al., 2004), could potentially be investigated as correlates of children's experientially driven conceptual advances. It is interesting to note that while developmental changes in TPJ selectivity are problematic for modularity theories, they are not problematic for theory theory, as one would expect changes in neural structures across childhood in response to theory revision and conceptual advances in ToM understanding.

Because only a few studies address the question of what conceptual change in ToM might look like at the neural level, researchers may draw inspiration from other neuroimaging fields in which the relation between conceptual and neural change has been more fully mapped. One such case involves the differential neural activations that subserve implicit and explicit memory. Event-related potentials have shown different spatiotemporal components of explicit and implicit memory retrieval (e.g., Paller et al., 2003; Rugg et al., 1998). In addition, Schott et al. (2005) found that implicit memory retrieval relied on prefrontal, fusiform, and extrastriate regions whereas explicit memory retrieval recruited posterior cingulate, precuneus,

and inferior parietal lobule. This example may be particularly relevant to theory theory in that the fledgling forms of mentalizing evident in infancy and very early childhood have been argued to represent an implicit ToM, whereas more verbal forms of ToM evident in preschoolers' false belief performance may represent an explicit ToM (e.g., Apperly and Butterfill, 2009; Clements and Perner, 1994; Clements et al., 2000; Low and Perner, 2012; Thoermer et al., 2012; Ruffman et al., 2001; Wang et al., 2012). To the extent that this is so, implicit and explicit processes supporting ToM may rely on partially independent brain networks just as implicit and explicit memory processes do. Further, the change from implicit to explicit ToM may occur because of conceptual advances in the understanding of mental state concepts. Consistent with that possibility, Wellman et al. (2008) showed that infants' attention to intentional action was predictive of their ToM understanding as preschoolers, suggesting a link from implicit to explicit understanding. Clearly, however, for neural evidence to support theory theory we would need to determine (a) whether different neural signatures indeed underlie implicit versus explicit ToM, and (b) whether neural changes occur during the transition from implicit to explicit ToM understanding, specifically ones which take the form of building connections between brain regions that support these two types of ToM reasoning.

In assessing the neural basis of theory change it would also be worth examining the role of the default mode network, a coordinated network that is activated by task-independent introspection or self-referential thought (e.g., Fair et al., 2008; Sheline et al., 2009) and is especially active during high-level social cognitive tasks (Harrison et al., 2008). This network involves several neural regions overlapping with those discussed in this review, such as MPFC, medial precuneus/posterior cingulate, and TPJ (Supekar et al., 2010). Functional connectivity in the default mode network is not yet fully mature in childhood (7–9-year-old); its regions only become fully connected into a cohesive network by early adulthood (Fair et al., 2008), consistent with predictions of theory theory that conceptual change follows an extended trajectory over development. We suggest that a network-wide approach may be the optimal way to examine the predictions of theory theory, as conceptual changes plausibly involve a broader set of regions and processes compared to those involved in the other theories that tend to target more specific regions or networks associated with modular ToM processing, simulation, or executive processes.

To summarize, the neural evidence for *theory change* in ToM is fairly sparse, although recent work is beginning to address this theory (Bowman et al., 2012; Sabbagh, 2011). Clearly, in future work very young children will need to be assessed to determine whether a particular neural system supports advances in mental state understanding through hypothesis testing. Further, the development of these neural systems should correlate with behavioral performance on ToM tasks. Research of this kind would ideally take a longitudinal approach starting with young children who do not yet possess the relevant ToM concepts and following them through the transition to concept acquisition.

6. Conclusions

6.1. Challenges

Our review illustrates that social-cognitive neuroscience research has contributed importantly to our understanding of the neural systems that support ToM. Each of the theories discussed is supported by some neural evidence, yet each is also called into question by other evidence or left open to question in the absence of sufficient evidence. Moving forward, different tactics may need to be utilized to better interrogate the theories. As we see it, at least five issues currently pose challenges to research examining the neural evidence for theories of ToM acquisition. Although these challenges may seem daunting, we believe that advances in current technologies and methodologies, as well as better theoretical articulation of core conceptual issues, may result in significant forward progress.

A first difficulty is that theories of ToM acquisition need to be sharpened so that neuroimaging research can better distinguish among them rather than simply amass evidence in support of them. Part of the problem is that the theories are not necessarily entirely mutually exclusive. For example, Leslie et al.'s modularity account explicitly includes an ancillary executive component (the selection processor). Similarly, there is no reason in principle that simulation theory or theory theory could not have at least *some* modular basis that provides input into the more flexible, general-purpose mechanisms postulated by these theories. In addition, there is no reason why simulation could not be a useful tool for a child theoretician (just as it is for a scientific theoretician). Finally, developments in executive function could, and likely do, play a role in facilitating more accurate simulation and more sophisticated theory building. And yet the theories are sufficiently vague on these points that it is often difficult to assess the extent to which evidence in support of one theory is damaging to the other theories. Until the theories are better articulated in these and related ways neural evidence is unlikely to fully discriminate among them (and, of course, the same is true for behavioral evidence).

A second, related challenge is that several of the proposed mechanisms behind ToM acquisition are difficult to articulate sufficiently to be easily tested at the neural level. For example, conceptual change, the central mechanism in theory theory, is a rather abstract construct that could take many forms at both the behavioral and neural levels (although see Gopnik and Wellman, 2012, for an attempt to translate theory change into more precise computational form). Similarly, simulation is difficult to define, as there are several routes by which the contents of another's mind can be approximated. Moreover, many questions remain concerning simulation such as the extent to which it is automatic/implicit or conscious/explicit. Such distinctions have implications for how simulation is measured. In contrast, with respect to modularity theory, progress in defining and quantifying neural selectivity (Saxe, 2010; Saxe et al., 2009), offers an encouraging example of how complex neural processes can be measured and used to test a theory. If we could similarly quantify conceptual change

or simulation in terms of expected differences in brain activation or increasing selectivity of neural structures, we would be better able to acquire evidence to support or refute the relevant theories.

A third challenge is addressing task heterogeneity within the ToM literature. Across the studies reviewed here, ToM paradigms varied enormously, although most focused on cognitive ToM tasks rather than affective ones. At least some of the differences and inconsistencies across studies could well be accounted for by varying task structures and task demands, as well as by varying conceptual content. Hence, the implications of these imaging studies for theories of ToM acquisition are not always straightforward. The same challenge arises from heterogeneity in age groups studied and heterogeneity in neuroimaging methods and analyses.

A fourth challenge is how to conceptualize *maturity* of the neural networks involved in ToM processing. Brain regions involved in ToM reasoning continue to develop into adolescence, as does behavioral ToM performance (Dumontheil et al., 2010), so it is important to clarify what criteria are being used to consider neural processing as 'mature' (see Pfeifer and Allen, 2012). For example, one could define maturity in terms of neural functioning that supports successful behavioral performance. So, even though 5-year-olds may not have fully developed brain regions in the ToM network, the neural processes that allow them to pass simple false belief tasks may be considered 'mature'. Alternately, maturity may be defined by the speed or efficiency of the neural systems that support ToM, in which case the neural systems underlying false belief understanding would not be considered mature until much later. Within the executive account, this issue is particularly salient as maturity of EF is predicted to influence ToM maturity. Here again, we know that (neurally as well as behaviorally) EF develops into early adulthood (Luna et al., 2010). However, the expression account predicts that at a certain level of EF, ToM should no longer be hindered by executive difficulties with the task, thereby enabling it to be expressed. As a result, it is just as important to define what counts as neurally mature EF for a particular purpose as it is to define neurally mature ToM for that purpose. Resolving this issue ideally would produce clear and consistent definitions of maturity that can be used in future neuroimaging studies. Moreover, what counts as maturity would need to be established separately for different ToM concepts, which of course each have their own developmental trajectories.

A final challenge is that ToM neuroimaging research consists predominantly of studies of online ToM processing in adults, while only a fraction of studies examine how ToM is acquired during development. Adult studies are certainly important in understanding how the fully developed form of ToM is characterized at the neural level, but until we have a clear developmental picture of how neural systems support the acquisition of ToM we will not be able to strongly test the different theoretical stances behind acquisition. There are two problems. First, the acquisition of more nuanced concepts by "experts" may be quite different from acquisition of initial basic concepts by "novices". Second, and more generally, it is not clear how the neural underpinnings of online ToM processing relate to the neural underpinnings of ToM acquisition. For example, if it is a matter of the magnitude of neural activation, with higher levels indicating mature processing and lower levels indicating ToM acquisition in progress (or vice versa), then acquisition and online processing may be viewed as a continuum in terms of neural activation. If these two modes of activation rely on different brain regions or systems, however, this would indicate that acquisition and online processing are relatively independent.

Studies attempting to measure the neural correlates of conceptual change would do well to focus on children who are on the cusp of acquiring the relevant ToM concept, and to monitor their transition to successful concept acquisition. Fortunately, technological and methodological advances increasingly allow very young children to be scanned while minimizing data loss (e.g., Cantlon et al., 2006; Gaffrey et al., 2011; Raschle et al., 2012), providing the opportunity to finally examine ToM acquisition as it unfolds. That said, successfully neuroimaging young children is not without challenges. For example, in contrast to behavioral measures of ToM such as false belief tasks that rely on a very small number of trials, imaging studies typically demand high repetition of trials. Keeping children attentive with engaging and varied tasks is clearly necessary but not always easy.

6.2. Future directions

Having now identified some of the major challenges that the field needs to address, we next suggest promising future directions that may help to clarify the neural basis of ToM acquisition. First, we know that both structural and functional changes occur in neural regions associated with social cognitive development. For example, Mills et al. (2012) found that gray matter and cortical thickness decrease from childhood into early adulthood in mentalizing regions such as MPFC and TPJ, and that the surface areas of these regions peak in early or pre-adolescence before decreasing in the early twenties. These structural changes may have implications for development of mentalizing functions. Functionally, there are hints of declines in MPFC activation during mentalizing from early adolescence to adulthood (Gunther Moor et al., 2012) as well as a potential shift from anterior to posterior brain regions supporting mentalizing (e.g., Blakemore, 2012; Burnett et al., 2009; Gunther Moor et al., 2012). Yet, few studies have simultaneously examined structural and functional changes in brain areas associated with ToM development (see Gunther Moor et al., 2012) and fewer still have investigated young children. Doing so will be an important next step in this area of research.

Second, training and/or microgenetic studies investigating neural changes that accompany ToM development would greatly add to our knowledge. For example, what is the impact of training EF, simulation, or ToM reasoning itself on the neural regions associated with mentalizing? This is perhaps one of the most promising avenues of work for differentiating among theories, as the theories predict that different types of training will be effective and should generate distinctive neural changes in some of the brain regions discussed earlier. Interestingly, using

neuroimaging techniques to examine the impact of ToM training on the neural structures of individuals with autism is already underway (Bolte et al., 2006). On a practical note, understanding what processes and neural regions are driving ToM development has implications for intervention and training programs. For example, children's simulation or executive abilities may be able to be trained in order to improve ToM, whereas, if neural evidence suggests that an early developed module is responsible for ToM reasoning then training these abilities later in childhood may be unproductive.

Finally, using TMS to examine the brain areas that are associated with simulation, conceptual advances, or EF may illuminate the role these processes play in ToM development (e.g., Young et al., 2010b). Disrupting key ToM regions may result in immature behavioral performance, making adults' performance child-like on certain ToM tasks, and so helping to establish the neural regions responsible for early ToM development. In addition, studies that temporarily disrupt neural regions associated with a certain type of processing (whether simulation, executive, or conceptual) may indicate whether these processes work independently or simultaneously to contribute to mentalizing. For example, disrupting the rTPJ may result in poorer ToM performance, but are there other neural regions such as MPFC or pSTS that account for individual differences in ToM performance independent of the rTPJ? In other words, can one succeed on ToM tasks without the TPJ?

6.3. Conclusion

The current state of neuroimaging research has important implications for theories of ToM acquisition. Although current evidence does not allow us to fully distinguish among theories, significant advances have been made in examining each theory. First, the lack of consistency of recruitment of the TPJ across development and lack of TPJ specificity challenges the idea that TPJ could be a dedicated ToMM of the kind proposed in certain strict varieties of modularity theory (Leslie et al., 2004). Second, both CMS and MNS have been identified as plausible candidate mechanisms for simulation theories, although it is unclear how these systems contribute to ToM acquisition. Third, evidence for the role of executive processes in ToM suggests that inhibition is necessary, but not sufficient for ToM processing, as several neural regions seem to contribute to mentalizing in addition to regions supporting inhibitory processing. Finally, evidence that the TPJ may be involved in resolving discrepancies between internal expectations and external reality suggests that TPJ may play a role in theory revision and conceptual change.

It is our hope that with further advances in the field such as the ability to conduct structural and functional MRI with even younger children, it will become easier to tease theories of ToM acquisition apart with neural evidence. The future directions we have outlined represent promising avenues of research that should generate a richer appreciation of *how* ToM develops in the brain, building upon past work that has described *where* the neural substrates of ToM are in the brain. We hope this review illustrates key neural regions and processes that might play a role in

ToM development and how advances in neuroimaging may offer fruitful ways in which to examine the development of ToM and, as a result, to address the theories behind its acquisition.

Conflict of interest

None declared.

Acknowledgment

References

Ames, D.R., 2004a. Inside the mind reader's tool kit: projection and stereotyping in mental state inference. Journal of Personality and Social Psychology 87, 340–353.

Ames, D.R., 2004b. Strategies for social inference: a similarity contingency model of projection and stereotyping in attribute prevalence estimates. Journal of Personality and Social Psychology 87, 573–585.

Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. Nature Reviews Neuroscience 7, 268–277.

Apperly, I.A., 2008. Beyond simulation-theory and theory-theory: why social cognitive neuroscience should use its own concepts to study theory of mind. Cognition 107, 266–283.

Apperly, I.A., Butterfill, S.A., 2009. Do humans have two systems to track beliefs and belief-like states? Psychological Review 116, 953–970.

Apperly, I.A., Samson, D., Chiavarino, C., Humphreys, G.W., 2004. Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. Journal of Cognitive Neuroscience 16, 1773–1784.

Apperly, I.A., Samson, D., Humphreys, G.W., 2009. Studies of adults can inform accounts of theory of mind development. Developmental Psychology 45, 190–201.

Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. Trends in Cognitive Sciences 8, 170–177.

Arzy, S., Thut, G., Mohr, C., Michel, C.M., Blanke, O., 2006. Neural basis of embodiment: distinct contributions of temporoparietal junction and extrastriate body area. Journal of Neuroscience 26, 8074–8081.

Astington, J.W.E., Baird, J.A., 2005. Why language matters for theory of mind. In: Why Language Matters for Theory of Mind, Apr. 2002, University of Toronto, Toronto, ON, Canada; This Volume Originated from the Aforementioned Conference, Oxford University Press.

Baillargeon, R., Scott, R.M., He, Z., 2010. False belief understanding in infants. Trends in Cognitive Sciences 14, 110–118.

Baldwin, D.A., Baird, J.A., Saylor, M.M., Clark, M.A., 2001. Infants parse dynamic action. Child Development 72, 708–717.

Baldwin, D.A., Moses, L.J., 1994. The mindreading engine: evaluating the evidence for modularity. Current Psychology of Cognition 13, 553–560.

Baron-Cohen, S., 1994. How to build a baby that can read minds: cognitive mechanisms in mindreading. Current Psychology of Cognition 13, 513–552.

Baron-Cohen, S., 1995. Mind Blindness: An Essay on Autism and Theory of Mind. MIT Press, Cambridge, MA.

Baron-Cohen, S., 1998. Does the study of autism justify minimalist innate modularity? Cognitive Development 10, 179–191.

Behne, T., Carpenter, M., Call, J., Tomasello, M., 2005. Unwilling versus unable: infants' understanding of intentional action. Developmental Psychology 41, 328–337.

Blakemore, S.J., 2012. Imaging brain development: the adolescent brain. NeuroImage 61, 397–406.

Bolte, S., Hubl, D., Feineis-Matthews, S., Prvulovic, D., Dierks, T., Poustka, F., 2006. Facial affect recognition training in autism: can we animate the fusiform gyrus? Behavioral Neuroscience 120, 211–216.

Bowman, L.C., Liu, D., Meltzoff, A.N., Wellman, H.M., 2012. Neural correlates of belief-and desire-reasoning in 7-and 8-year-old children: an event-related potential study. Developmental Science 15, 618–632.

Bunge, S.A., Dudukovic, N.M., Thomason, M.E., Vaidya, C.J., Gabrieli, J.D.E., 2002. Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. Neuron 33, 301–311.

Burnett, S., Bird, G., Moll, J., Frith, C., Blakemore, S.-J., 2009. Development during adolescence of the neural processing of social emotion. Journal of Cognitive Neuroscience 21, 1736–1750.

Cantlon, J.F., Brannon, E.M., Carter, E.J., Pelphrey, K.A., 2006. Functional imaging of numerical processing in adults and 4-year-old children. PLoS Biology 4, e125.

Carlson, S.M., Claxton, L.J., Moses, L.J., 2014. The relation between executive function and theory of mind is more than skin deep. Journal of Cognition and Development (in press).

Carlson, S.M., Moses, L.J., 2001. Individual differences in inhibitory control and children's theory of mind. Child Development 72, 1032–1053.

Carlson, S.M., Moses, L.J., Breton, C., 2002. How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. Infant and Child Development 11, 73–92.

Carlson, S.M., Moses, L.J., Claxton, L.J., 2004. Individual differences in executive functioning and theory of mind: an investigation of inhibitory control and planning ability. Journal of Experimental Child Psychology 87, 299–319.

Carlson, S.M., Moses, L.J., Hix, H., 1998. The role of inhibitory processes in young children's difficulties with deception and false belief. Child Development 69, 672–691.

Carpendale, J.I.M., Lewis, C., 2004. Constructing an understanding of mind. The development of children's social understanding within social interaction. Behavioral and Brain Sciences 27, 79–96.

Carruthers, P., 2003. Moderately massive modularity. Royal Institute of Philosophy Supplement, 67–90.

Casey, B.J., Trainor, R.J., Orendi, J.L., Schubert, A.B., Nystrom, L.E., Giedd, J.N., et al., 1997. A developmental functional MRI study of prefrontal activation during performance of a go-nogo task. Journal of Cognitive Neuroscience 9, 835–847.

Chandler, M.J., 1973. Egocentrism and antisocial behavior: the assessment and training of social perspective-taking skills. Developmental Psychology 9, 326–332.

Clements, W.A., Rustin, C.L., McCallum, S., 2000. Promoting the transition from implicit to explicit understanding: a training study of false belief. Developmental Science 3, 81–92.

Clements, W.A., Perner, J., 1994. Implicit understanding of belief. Cognitive Development 9, 377–395.

Coltheart, M., 1999. Modularity and cognition. Trends in Cognitive Sciences 3, 115–120.

Corbetta, M., Patel, G., Shulman, G.L., 2008. The reorienting system of the human brain: from environment to theory of mind. Neuron 58, 306–324.

Cutting, A.L., Dunn, J., 1999. Theory of mind, emotion understanding, language, and family background: individual differences and interrelations. Child Development 70, 853–865.

Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y., Iacoboni, M., 2006. Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. Nature Neuroscience 9, 28–30.

Denny, B., Kober, H., Wager, T.D., Ochsner, K.N., 2012. A meta-analysis of functional neuroimaging studies of self and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. Journal of Cognitive Neuroscience 24, 1742–1752.

Dohnel, K., Schuwerk, T., Meinhardt, J., Sodian, B., Hajak, G., Sommer, M., 2012. Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. NeuroImage 60, 1652–1661.

Dumontheil, I., Apperly, I.A., Blakemore, S.-J., 2010. Online usage of theory of mind continues to develop in late adolescence. Developmental Science 13, 331–338.

Durston, S., Casey, B.J., 2006. What have we learned about cognitive development from neuroimaging? Neuropsychologia 44, 2149–2157.

Durston, S., Thomas, K.M., Yang, Y., Ulug, A.M., Zimmerman, R.D., Casey, B.J., 2002. A neural basis for the development of inhibitory control. Developmental Science 5, F9–F16.

Fair, D.A., Cohen, A.L., Dosenbach, N.U., Church, J.A., Miezin, F.M., Barch, D.M.. . ., Schlaggar, B.L., 2008. The maturing architecture of the brain's default network. Proceedings of the National Academy of Science 105, 4028–4032.

Fanning, J.R., Bell, M.D., Fiszdon, J.M., 2012. Is it possible to have impaired neurocognition but good social cognition in schizophrenia? Schizophrenia Research 135, 68–71.

Fassbender, C., Simones-Franklin, C., Murphy, K., Hester, R., Meaney, J., Robertson, I.H., Gravan, H., 2006. The role of a right fronto-parietal network in cognitive control: common activations for cues-to-attend and response inhibition. Journal of Psychophysiology 20, 286–296.

Fodor, J., 1983. Modularity of Mind. MIT Press, Cambridge, MA.

Friston, K.J., Buechel, C., Fink, G.R., Morris, J., Rolls, E., Dolan, R.J., 1997. Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6, 218–229.

Gaffrey, M.S., Luby, J.L., Belden, A.C., Hirshberg, J.S., Volsch, J., Barch, D.M., 2011. Association between depression severity and amygdala reactivity during sad face viewing in depressed preschoolers: an fMRI study. Journal of Affective Disorders 129, 364–370.

Gallese, V., 2007. Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. Philosophical Transactions of the Royal Society B: Biological Sciences 362, 659–669.

Gallese, V., Goldman, A., 1998. Mirror neurons and the simulation theory of mind-reading. Trends in Cognitive Sciences 2, 493–501.

Geng, J.J., Mangun, G.R., 2011. Right temporoparietal junction activation by a salient contextual cue facilitates target discrimination. NeuroImage 54, 594–601.

Geng, J.J., Vossel, S., 2013. Re-evaluating the role of TPJ in attentional control: contextual updating? Neuroscience and Biobehavioral Reviews 37, 2608–2620.

German, T.P., Hehman, J.A., 2006. Representational and executive selection resources in 'theory of mind': evidence from compromised belief-desire reasoning in old age. Cognition 101, 129–152.

Glenberg, A.M., 2011. Introduction to the mirror neuron forum. Perspectives on Psychological Science 6, 363–368.

Goldman, A., 2009. Mirroring, simulating, and mindreading. Mind and Language 24, 235–252.

Gopnik, A., 2003. The theory theory as an alternative to the innateness hypothesis. In: Antony, L., Hornstein, N. (Eds.), Chomsky and his Critics. Basil Blackwell, New York.

Gopnik, A., Wellman, H.M., 1994. The theory theory. In: Hirschfeld, L., Gelman, S. (Eds.), Mapping the Mind. Cambridge University Press, New York, pp. 257–293.

Gopnik, A., Wellman, H.M., 2012. Reconstructing constructivism: causal models, Bayesian learning mechanisms, and the theory theory. Psychological Bulletin 138, 1085–1108.

Gordon, R.M., 1992. The simulation theory: objections and misconceptions. Mind and Language 7, 11–34.

Gunther Moor, B., de Macks, Z.A.O., Güroğlu, B., Rombouts, S.A., Van der Molen, M.W., Crone, E.A., 2012. Neurodevelopmental changes of reading the mind in the eyes. Social Cognitive and Affective Neuroscience 7, 44–52.

Gweon, H., Dodell-Feder, D., Bedny, M., Saxe, R., 2013. Theory of mind performance in children correlates with functional specialization of a brain region for thinking about thoughts. Child Development 83, 1853–1868.

Happé, F., Frith, U., 2013. Annual research review: towards a developmental neuroscience of atypical social cognition. Journal of Child Psychology and Psychiatry (in press).

Harris, P.L., 2000. The Work of the Imagination. Blackwell, Oxford, England.

Harris, P.L., 1992. From simulation to folk psychology: the case for development. Mind and Language 7, 120–144.

Harrison, B.J., Pujol, J., López-Solà, M., Hernández-Ribas, R., Deus, J., Ortiz, H.. . ., Cardoner, N., 2008. Consistency and functional specialization in the default mode brain network. Proceedings of the National Academy of Science 105, 9781–9786.

Hartwright, C.E., Apperly, I.A., Hansen, P.C., 2012. Multiple roles for executive control in belief–desire reasoning: Distinct neural networks are recruited for self perspective inhibition and complexity of reasoning. NeuroImage 61, 921–930.

Hughes, C., 1998. Executive function in preschoolers: links with theory of mind and verbal ability. British Journal of Developmental Psychology 16, 233–253.

Hughes, C., Leekham, 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, 590–619.

Iacoboni, M., 2009. Imitation, empathy, and mirror neurons. Annual Review of Psychology 60, 653–670.

Iacoboni, M., Dapretto, M., 2006. The mirror neuron system and the consequences of its dysfunction. Nature Reviews Neuroscience 7, 942–951.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., Rizzolatti, G., 2005. Grasping the intentions of others with one's own mirror neuron system. PLoS Biology 3, e79.

Jenkins, J.M., Astington, J.W., 1996. Cognitive factors and family structure associated with theory of mind development in young children. Developmental Psychology 32, 70–78.

Jenmalm, P., Schmitz, C., Forssberg, H., Ehrsson, H.H., 2006. Lighter or heavier than predicted: neural correlates of corrective mechanisms during erroneously programmed lifts. Journal of Neuroscience 26, 9015–9021.

Keller, T.A., Just, M.A., 2009. Altering cortical connectivity: remediation-induced changes in the white matter of poor readers. Neuron 64, 624–631.

Keysar, B., Lin, S., Barr, D.J., 2003. Limits on theory of mind use in adults. Cognition 89, 25–41.

Kobayashi, C., Glover, G., Temple, C., 2007. Children's and adults' neural bases of verbal and nonverbal 'Theory of Mind'. Neuropsychologia 45, 1522–1532.

Kovacs, A.M., Teglas, E., Endress, A.D., 2010. The social sense: susceptibility to others' beliefs in human infants and adults. Science 330, 1830–1834.

Kubit, B., Jack, A.I., 2014. Rethinking the role of the rTPJ in attention and social cognition in light of the opposing domains hypothesis: findings from an ALE-based meta-analysis and resting-state functional connectivity. Frontiers in Human Neuroscience 7, 323.

Lenggenhager, B., Tadi, T., Metzinger, T., Blanke, O., 2007. Video ergo sum: manipulating bodily self-consciousness. Science 317, 1096–1099.

Leslie, A.M., Friedman, O., German, T.P., 2004. Core mechanisms in 'theory of mind'. Trends in Cognitive Sciences 8, 528–533.

Leslie, A.M., Polizzi, P., 1998. Inhibitory processing in the false-belief task: two conjectures. Developmental Science 1, 247–254.

Levy, B.J., Wagner, A.D., 2011. Cognitive control and right ventrolateral prefrontal cortex: reflexive reorienting, motor inhibition, and action updating. Annals of the New York Academy of Sciences 1224, 40–62.

Lillard, A.S., Kavanaugh, R.L., 2014. A longitudinal study of the relationship between symbolic understanding and theory of mind in preschoolers. Child Development (in press).

Liu, D., Meltzoff, A.N., Wellman, H.M., 2009. Neural correlates of belief- and desire-reasoning. Child Development 80, 1163–1171.

Lohmann, H., Tomasello, M., 2003. The role of language in the development of false belief understanding: a training study. Child Development 74, 1130–1144.

Low, J., Perner, J., 2012. Implicit and explicit theory of mind: state of the art. British Journal of Developmental Psychology 30, 1–13.

Luna, B., Padmanabhan, A., O'Hearn, K., 2010. What has fMRI told us about the development of cognitive control through adolescence? Brain and Cognition 72, 101–113.

Mabbott, D.J., Noseworthy, M., Bouffet, E., Laughlin, S., Rockel, C., 2006. White matter growth as a mechanism of cognitive development in children. NeuroImage 33, 936–946.

Marshall, P.J., Meltzoff, A.N., 2011. Neural mirroring systems: exploring the EEG mu rhythm in human infancy. Developmental Cognitive Neuroscience 1, 110–123.

Marshall, P.J., Young, T., Meltzoff, A.N., 2010. Neural correlates of action observation and execution in 14-month-old infants: an event-related EEG desynchronization study. Developmental Science 14, 474–480.

Meltzoff, A.N., 2007. The 'like me' framework for recognizing and becoming an intentional agent. Acta Psychologica 124, 26–43.

Mills, K.L., Lalonde, F., Clasen, L.S., Giedd, J.N., Blakemore, S.J., 2012. Developmental changes in the structure of the social brain in late childhood and adolescence. Social Cognitive and Affective Neuroscience (in press).

Mitchell, J.P., 2008. Activity in right temporo-parietal junction is not selective for theory-of-mind. Cerebral Cortex 18, 262–271.

Mitchell, J.P., Banaji, M.R., Macrae, C.N., 2005. The link between social cognition and self-referential thought in the medial prefrontal cortex. Journal of Cognitive Neuroscience 17, 1306–1315.

Mitchell, J.P., Macrae, C.N., Banaji, M.R., 2006. Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. Neuron 50, 1–9.

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

Moses, L.J., Tahiroglu, D., 2010. Clarifying the relation between executive function and children's theories of mind. In: Carpendale, J., Iarocci, G., Müller, U., Sokol, B., Young, A. (Eds.), Self- and Social Regulation: Exploring the Relations between Social Interaction, Social Cognition, and the Development of Executive Functions. Oxford University Press, Oxford, pp. 218–231.

Nagy, Z., Westerberg, H., Klingberg, T., 2004. Maturation of white matter is associated with the development of cognitive functions during childhood. Journal of Cognitive Neuroscience 16, 1227–1233.

Nichols, S., Stich, S., Leslie, A., Klein, D., 1995. Varieties of off-line simulation. In: Carruthers, P., Smith, P. (Eds.), Theories of Theories of Mind. Cambridge University Press, Cambridge, pp. 39–74.

Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.-B., 2005. Valid conjunction inference with the minimum statistic. NeuroImage 25.

Ochsner, K.N., Beer, J.S., Robertson, E.R., Cooper, J.C., Gabrieli, J.D., Kihlstrom, J.F., et al., 2005. The neural correlates of direct and reflected self-knowledge. NeuroImage 28, 797–814.

Onishi, K.H., Baillargeon, R., 2005. Do 15-month-old infants understand false belief? Science 308, 255–258.

Ozonoff, S., Miller, J.N., 1995. Teaching theory of mind: a new approach to social skills training for individuals with autism. Journal of Autism and Developmental Disorders 25, 415–433.

Paller, K.A., Hutson, C.A., Miller, B.B., Boehm, S.G., 2003. Neural manifestations of memory with and without awareness. Neuron 38, 507–516.

Pears, K.C., Moses, L.J., 2003. Demographics, parenting, and theory of mind in preschool children. Social Development 12, 1–20.

Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., Ladurner, G., 2006. Thinking of mental and other representations: the roles of left and right temporo-parietal junction. Social Neuroscience 1, 245–258.

Pfeifer, J.H., Allen, N., 2012. Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. Trends in Cognitive Sciences 16, 322–329.

Pfeifer, J.H., Iacoboni, M., Mazziotta, J.C., Dapretto, M., 2008. Mirroring others' emotions elates to empathy and social abilities during childhood. NeuroImage 39, 2076–2085.

Pfeifer, J.H., Lieberman, M.D., Dapretto, M., 2007. "I know you are but what am I?" Neural bases of self and social knowledge retrieval in children and adults. Journal of Cognitive Neuroscience 19, 1323–1337.

Pfeifer, J.H., Masten, C.L., Borofsky, L.A., Dapretto, M., Fuligni, A.J., Lieberman, M.D., 2009. Neural correlates of direct and reflected self-appraisals in adolescents and adults: when social perspective taking informs self-perception. Child Development 80, 1016–1038.

Pfeifer, J.H., Merchant, J.S., Colich, N.L., Rudie, J.D., Dapretto, M., 2011. Self-other differentiation in cortical midline structures is atypical in children and adolescents with high-functioning ASD. In: Paper session presented at the International Meeting for Autism Research, San Diego, CA, May 2011.

Pfeifer, J.H., Merchant, J.S., Colich, N., Hernandez, L., Rudie, J., Dapretto, M., 2013. Neural and behavioral responses during self-evaluative processes differ in youth with and without autism. Journal of Autism and Developmental Disorders 43, 272–285.

Phillips, A.T., Wellman, H.M., Spelke, E.S., 2002. Infants' ability to connect gaze and emotional expression to intentional action. Cognition 85, 53–78.

Raschle, N.M., Zuk, J., Gaab, N., 2012. Functional characteristics of developmental dyslexia in left-hemispheric posterior brain regions predate reading onset. Proceedings of the National Academy of Science 109, 2156–2161.

Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. Annual Review of Neuroscience 27, 169–192.

Rothmayr, C., Sodian, B., Hajak, G., Dohnel, K., Meinhardt, J., Sommer, M., 2010. Common and distinct neural networks for false-belief reasoning and inhibitory control. NeuroImage 56, 1705–1713.

Ruffman, T., Garnham, W., Import, A., Connolly, D., 2001. Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. Journal of Experimental Child Psychology 80, 201–224.

Ruffman, T., Slade, L., Crowe, E., 2002. The relation between children's and mothers' mental state language and theory-of-mind understanding. Child Development 73, 734–751.

Rugg, M.D., Mark, R.E., Walla, P., Schloerscheidt, A.M., Birch, C.S., Allan, K., 1998. Dissociation of the neural correlates of implicit and explicit memory. Nature 392, 595–598.

Russell, J., 1997. How executive disorders can bring about an inadequate 'theory of mind'. In: Russell, J. (Ed.), Autism as an Executive Disorder. Oxford University Press, New York, NY, pp. 256–304.

Sabbagh, M.A., 2011. Adjusting expectations: understanding the role of the right temporal–parietal juncture in representational theory of mind. In: Paper session presented at the Society for Research in Child Development Meeting Montreal, QC, Canada, April 2011.

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

Sabbagh, M.A., Hopkins, S.F.R., Benson, J.E., Flanagan, J.R., 2010. Conceptual change and preschoolers' theory of mind: evidence from load–force adaptation. Neural Networks 23, 1043–1050.

Sabbagh, M.A., Moses, L.J., Shiverick, S., 2006. Executive functioning and preschoolers' understanding of false beliefs, false photographs, and false signs. Child Development 77, 1034–1049.

Samson, D., 2009. Reading other people's mind: Insights from neuropsychology. Journal of Neuropsychology 3, 3–16.

Samson, D., Apperley, I.A., Chiavarino, C., Humphreys, G.W., 2004. Left temporoparietal junction is necessary for representing someone else's beliefs. Nature Neuroscience 7, 499–500.

Saxe, R., 2010. The right temporo-parietal junction: a specific brain region for thinking about thoughts. In: Leslie, A., German, T. (Eds.), Handbook of Theory of Mind.

Saxe, R., 2006. Why and how to study theory of mind with fMRI. Brain Research 1079, 57–65.

Saxe, R., 2009. The neural evidence for simulation is weaker than I think you think it is. Philosophical Studies 144, 447–456.

Saxe, R., Kanwisher, N., 2003. People thinking about thinking people. The role of the temporo-parietal junction in theory of mind. NeuroImage 19, 1835–1842.

Saxe, R., Moran, J.M., Scholz, J., Gabrieli, J., 2006. Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. Social Cognitive and Affective Neuroscience 1, 229–234.

Saxe, R., Powell, L.J., 2006. It's the thought that counts: specific brain regions for one component of theory of mind. Psychological Science 17, 692–699.

Saxe, R., Wexler, A., 2005. Making sense of another mind: the role of the right temporo-parietal junction. Neuropsychologia 43, 1391–1399.

Saxe, R., Whitfield-Gabrieli, S., Scholz, J., Pelphrey, K.A., 2009. Brain regions for perceiving and reasoning about other people in school-aged children. Child Development 80, 1197–1209.

Saylor, M., Baldwin, D., Baird, J., LaBounty, J., 2007. Infants' on-line segmentation of dynamic human action. Journal of Cognition and Development 8, 113–128.

Scholl, B.J., Leslie, A.M., 2001. Minds, modules, and meta-analysis. Child Development 72, 696–701.

Scholl, B.J., Leslie, A.M., 1999. Modularity, development and 'theory of mind'. Mind and Language 14, 131–153.

Scholz, J., Triantafyllous, C., Whitfield-Gabrieli, S., Brown, E.N., Saxe, R., 2009. Distinct regions of the right temporo-parietal junction are selective for theory of mind and exogenous attention. PLoS ONE 4, 1–7.

Schott, B.H., Henson, R.N., Richardson-Klavehn, A., Becker, C., Thoma, V., Heinze, H.J., Düzel, E., 2005. Redefining implicit and explicit memory: the functional neuroanatomy of priming, remembering, and control of retrieval. Proceedings of the National Academy of Sciences of the United States of America 102, 1257–1262.

Schwanenflugel, P., Fabricius, W., Noyes, C., 1996. Developing organization of mental verbs: evidence for the development of a constructivist theory of mind in middle childhood. Cognitive Development 11, 265–294.

Sheline, Y.I., Barch, D.M., Price, J.L., Rundle, M.M., Vaishnavi, S.N., Snyder, A.Z., Raichle, M.E., 2009. The default mode network and self-referential processes in depression. Proceedings of the National Academy of Science 106, 1942–1947.

Slaughter, V., Gopnik, A., 1996. Conceptual coherence in the child's theory of mind: training children to understand belief. Child Development 67, 2967–2988.

Southgate, V., Johnson, M.H., El Karoui, I., Csibra, G., 2010. Motor system activation reveals infants' on-line prediction of others' goals. Psychological Science 21, 355–359.

Southgate, V., Johnson, M.H., Osborne, T., Csibra, G., 2009. Predictive motor activation during action observation in human infants. Biology Letters 5, 769–772.

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, 1112–1123.

Spreng, R.N., Mar, R.A., Kim, A.S.N., 2009. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. Journal of Cognitive Neuroscience 21, 489–510.

Stone, V.E., Baron-Cohen, S., Knight, R.T., 1998. Frontal lobe contributions to theory of mind. Journal of Cognitive Neuroscience 10, 640–656.

Supekar, K., Uddin, L.Q., Prater, K., Amin, H., Greicius, M.D., Menon, V., 2010. Development of functional and structural connectivity within the default mode network in young children. NeuroImage 52, 290–301.

Taylor, M., Carlson, S.M., 1997. The relation between individual differences in fantasy and theory of mind. Child Development 68, 436–455.

Thatcher, R.W., 1992. Cyclic cortical reorganization in early childhood. Brain and Cognition 20, 24–50.

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, 172–187.

Uddin, L.Q., Iacoboni, M., Lange, C., Keenan, J.P., 2007. The self and social cognition: the role of cortical midline structures and mirror neurons. Trends in Cognitive Sciences 11, 153–157.

van der Meer, L., Groenewold, N.A., Nolen, W.A., Pijnenborg, M., Aleman, A., 2011. Inhibit yourself and understand the other: Neural basis of distinct processes underlying theory of mind. NeuroImage 56, 2364–2374.

Wang, B., Low, J., Jing, Z., Qinghua, Q., 2012. Chinese preschoolers' implicit and explicit false-belief understanding. British Journal of Developmental Psychology 30, 123–140.

Waytz, A., Mitchell, J.P., 2011. Two mechanisms for simulating other minds: dissociations between mirroring and self-projection. Current Directions in Psychological Science 20, 197–200.

Wellman, H.M., Cross, D., Watson, J., 2001. Meta-analysis of theory-of-mind development. The truth about false belief. Child Development 72, 655–684.

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

Wellman, H.M., Lopez-Duran, S., LaBounty, J., Hamilton, B., 2008. Infant attention to intentional action predicts preschool theory of mind. Developmental Psychology 44, 618–623.

Wilkinson, M.R., Ball, L.J., 2012. Why studies of autism spectrum disorders have failed to resolve the theory theory versus simulation theory debate. Review of Philosophy and Psychology 3, 263–291.

Yamaguchi, S., Zheng, D., Oka, T., Bokura, H., 2008. The key locus of common response inhibition network for no-go and stop signals. Journal of Cognitive Neuroscience 20, 1434–1442.

Young, L., Dodell-Feder, D., Saxe, R., 2010a. What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. Neuropsychologia 48, 2658–2664.

Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., Saxe, R., 2010b. Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. Proceedings of the National Academy of Science 107, 6753–6758.