



Editorial: The neuroscience of theory of mind

Rebecca Saxe & Simon Baron-Cohen

To cite this article: Rebecca Saxe & Simon Baron-Cohen (2006) Editorial: The neuroscience of theory of mind, *Social Neuroscience*, 1:3-4, 1-9, DOI: [10.1080/17470910601117463](https://doi.org/10.1080/17470910601117463)

To link to this article: <http://dx.doi.org/10.1080/17470910601117463>



Published online: 24 Feb 2007.



Submit your article to this journal [↗](#)



Article views: 490



View related articles [↗](#)



Citing articles: 11 View citing articles [↗](#)

Editorial: The neuroscience of theory of mind

Rebecca Saxe

MIT and Harvard University, Cambridge, MA, USA

Simon Baron-Cohen

University of Cambridge, Cambridge, UK

INTRODUCTION

Why a Special Issue on the neuroscience of Theory of Mind (ToM)? Although ToM has been thoroughly investigated by developmental psychologists over the last 25 years, since the landmark publication of the paper by Wimmer and Perner (1983), the neuroscience of this important ability—arguably the defining difference between humans and all other species (Saxe, 2006)—has only recently begun to receive serious attention. It occurred to us as guest editors, and to the journal editorial board, that this topic was ripe for a Special Issue of *Social Neuroscience*, not least because there are now a growing number of scientists interested in this topic producing high quality work, but also because the existing studies are often disparate, and there was an opportunity to invite world leaders in this field to present their work in one place, in the hope of drawing the key strands together. In this editorial we provide some background to this along with our own overview of the terrific collection of papers we are pleased to have received.

DOMAIN GENERAL OR DOMAIN SPECIFIC?

ToM has most often been tested using the well-known “false belief” task (Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). In one

standard version (the “object transfer” problem), young children watch an object being moved without someone else’s knowledge. In order to correctly predict or explain the character’s subsequent action, the child must pay attention to the character’s belief, and not just to the actual location of the object (Dennett, 1978).

Although the false belief task has been used in literally hundreds of studies, it remains controversial whether success on this task depends on the deployment of a “special” domain-specific mechanism for reasoning about other minds. Stone and Gerrans (this issue) provide a useful review of this debate. As they note, a domain-specific mechanism could not be *sufficient* for passing false belief tasks (Bloom & German, 2000). In order to recognize and reason about someone else’s false beliefs, children must use general cognitive abilities, including general perceptual and linguistic representations of the story, working memory (to track all the moving parts of the story), and motor representations of their own responses. The question is therefore whether in addition to *domain general* mechanisms, it is also necessary to postulate a distinct, dedicated *domain specific* mechanism for reasoning about beliefs and desires.

Traditionally, the debate about domain-specificity has considered evidence from typically developing children, and from children with autism (Baron-Cohen, Leslie, & Frith 1985; Baron-Cohen, Tager-Flusberg, & Cohen, 2000). In particular, emerging success on false belief

Correspondence should be addressed to: Rebecca Saxe, 46–4019, 43 Vassar St, MIT, Cambridge, MA 02138, USA.
E-mail: saxe@mit.edu

tasks may depend not on conceptual changes in ToM, but on children's ability to juggle two competing representations of reality (the actual state of affairs vs the world as represented in the character's head) and to inhibit an incorrect but compelling answer (the true location of the object). There is plenty of evidence that these components of inhibitory control both (a) develop around the same age as success on the false belief tasks and (b) are correlated with individual children's performance on false belief tasks (e.g., Carlson & Moses, 2001; but see Sabbagh et al., 2006, for intriguing recent evidence that differences in inhibitory control across cultures do not produce differences in false belief performance). Typically developing children's performance on false belief tasks closely matches performance on logically equivalent problems about non-mental false representations (Perner, Aichhorn, Kronbichler, Staffen, & Ladurner, this issue; Leslie & Thaiss, 1992; Zaitchik, 1990). It might be that what develops is therefore not a specific understanding of belief, but rather a general understanding of representational relationships—that is, a capacity for meta-representation. The clearest evidence of domain specificity to date has come from studies of autism. Children with autism show consistent delays or deficits in passing false belief tests (see Baron-Cohen et al., 2000, for a 15-year review), often in the face of intact performance on false photograph or false drawing tasks (Charman & Baron-Cohen, 1992; Leslie & Thaiss, 1992), suggesting not a general meta-representational problem but a domain-specific one as relates to representing mental states as *attitudes*.

Nevertheless, Stone and Gerrans (this issue) are not compelled by the evidence to date. These authors conclude that while false belief task performance does depend on domain-general mechanisms for inhibitory control and meta-representation, they see no reason yet to posit an additional domain-specific mechanism for belief attribution. Stone and Gerrans favor an argument from parsimony: scientists should not posit an extra mechanism if a smaller number of entities would in principle suffice. This theoretical position provides an excellent entrée for social neuroscientists. With neuroscientific methods, it is possible to get behind observed correlations in behavior, in order to directly test the hypothesis that reasoning about other minds depends on distinct neural mechanisms. Three papers in the

current issue use fMRI to investigate claims about domain-specificity.

Perner and colleagues (this issue) provide a direct response to Stone and Gerrans' (this issue) challenge. Both groups of authors agree that a mechanism that is specifically implicated in representing beliefs must meet at least three criteria: (1) a robust and replicable response to stories that do require belief attribution; (2) a significantly lower response to closely matched control stories that only require representing non-mental representations; and (3) no difference among non-mental stories that do or do not demand meta-representation. Perner and colleagues (this issue) provide evidence that one brain region meets all three of these criteria: the right temporo-parietal junction (RTPJ). First, they replicate the finding that the RTPJ is recruited when subjects read simple stories about characters' beliefs, but not for stories about physical representations like photographs (Saxe & Kanwisher, 2003). More importantly, the RTPJ is not recruited when subjects read stories about maps and signs, which are more similar to beliefs (and better correlated with belief-reasoning in development), because like beliefs, and unlike photographs, maps and signs are designed to represent the current reality and so can genuinely be false. Finally, the RTPJ did not distinguish between stories about false maps, that do depend on meta-representational understanding, and control stories about temporal change, that involve no meta-representation. These data thus provide a direct response to Stone and Gerrans' (this issue) challenge: the RTPJ does appear to reflect the functioning of a specialized, domain-specific mechanism for reasoning about beliefs. Also consistent with previous data, Perner and colleagues (this issue) found a different pattern in the left TPJ, which was recruited for both mental and non-mental meta-representational tasks (Apperly, Samson, Chiavarino, Bickerton, & Humphreys, 2006; Aichhorn, Perner, Kronbichler, Staffen, & Ladurner, 2006).

Further evidence that reasoning about beliefs depends on *both* domain-general and domain-specific mechanisms is presented in the paper by Saxe and colleagues (this issue). These authors first identify brain regions implicated in (a) belief reasoning and (b) domain-general response selection and inhibitory control. They then show that reasoning about beliefs does stimulate robust activity in domain-general mechanisms, but that an equal response is also observed when subjects

reason about false photographs. That is, domain-general mechanisms for inhibitory control, response selection, etc., are recruited by both false belief and false photograph stories, but there are additional brain regions that are recruited only for beliefs. These data suggest that belief reasoning depends on *both* domain-general and domain-specific mechanisms, while reading about false photographs depends only on the domain-general components (see Stone and Gerrans, Figure 1, p. 312). Next, Saxe and colleagues devised a non-verbal task on which subjects could view the same stimuli, and make the same responses, while either following domain-general rules or reasoning about beliefs. Consistent with Perner and colleagues' results (this issue), Saxe and colleagues found that the best evidence for a domain-specific response was observed in the right TPJ.

Saxe and colleagues also report that for the same stimuli, and the same responses, human adults were faster to make the response by reasoning about beliefs than by following domain-general rules. These results echo a whole tradition of research in evolutionary psychology. Cosmides & Tooby (1997), for example, showed that for logically identical “if–then” reasoning problems, most people are faster and more accurate if the problems are phrased in terms of social obligations than if they are phrased in general abstract terms. In the current issue, Ermer and colleagues use fMRI to support their central claim that humans have domain-specific mechanisms for social reasoning. They find brain regions recruited differentially when subjects solve the Wason selection task for social contracts (“If you borrow the car, then you must fill up the tank with gas”), compared to logically matched, equally familiar, precautionary rules (“If you work with TB patients, then you must wear a surgical mask”). While social contracts may not appear to involve ToM, a social contract assumes both parties to the contract *expect* the other to abide by the contract, and in this sense social contracts assume shared *beliefs* about the nature of the social obligation.

Across these papers, fMRI provides a useful method for determining the functional profile of brain regions, functions that are ultimately psychological. However, fMRI can only ever describe the brain regions that are *associated* with a psychological task; it does not describe what relationships are necessary. To address whether one cognitive function—e.g., language—is a necessary component of another—e.g., belief

attribution—social neuroscientists turn to other methods.

WHAT IS THE ROLE OF LANGUAGE IN TOM?

Mental states cannot be observed directly; beliefs and desires are invisible, abstract entities. Nor is there any simple correlation between mental states and behavior. For example, there is no observable behavior that is reliably diagnostic of whether a person currently believes that today is Tuesday (Baldwin & Saylor, 2005). One invaluable way to learn about the elusive contents of the mind is therefore to listen to how other people talk about the mind.

Research in developmental psychology suggests the importance of verbal communication for developing a theory of mind (Astington & Baird, 2005). Dunn and colleagues (e.g., Dunn & Brophy, 2005) first reported that language ability predicts success on the false belief task, independent of age. A similar correlation is observed in both samples of healthy children and of children with autism and other developmental disabilities (Astington & Jenkins, 1999; Peterson & Siegal, 1999). What is not clear is the direction of the relationship between communication and ToM. In all likelihood, it is a two-way relationship: Good language skills give you better access to another person's mental states (since language is a “print out” of someone else's mind), enabling you to learn more about their thoughts, intentions, and feelings. But ToM may also facilitate language-acquisition (Baron-Cohen, Baldwin, & Crowson, 1997; Bloom, 2000), for example, guiding the developing child to search for a speaker's intention in decoding novel words by following their gaze.

Clear evidence that reduced language input slows down the development of ToM comes from the finding that deaf children whose parents are non-native signers are selectively delayed in passing the false belief task (Peterson & Siegal, 1999). These children have difficulty even on non-verbal tests of false belief understanding, suggesting that the delay does not reflect the language demands of the tasks (e.g., Figueras-Costa & Harris, 2001). Moreover, even after accounting for the child's own language skills, the child's performance on the false belief task is independently predicted by the *mother's* proficiency with sign language, and specifically her use of mental

state words (Moeller & Schick, 2006). A recent investigation of Nicaraguan Sign Language suggests that performance on non-verbal false belief tasks is predicted by the age at which the child first learned sign language, and not by the number of years that they have been using it (Morgan & Kegl, 2006); children who first entered the sign language community before age 8 perform significantly better than those who did so later in life. Deaf children of native signers—who learn sign language from birth—show no delay at all (de Villiers, 2005). Clearly, non-verbal social information (available to all of the deaf populations) is not sufficient; linguistic exposure significantly influences the development of ToM.

In the current issue, Ronald and Plomin conducted a large-scale study of 9-year-old twins to investigate the origin of links between language and ToM. By comparing monozygotic and same-sex dizygotic twin pairs, the authors distinguish genetic influences on development from the effects of shared and non-shared environment. Ronald and Plomin found that ToM task performance in late childhood is moderately heritable, but also shows a significant effect of shared environment. The heritable aspect is important given the earlier mention of deficits in ToM in autism, since autism is a strongly genetic condition (Folstein & Rutter, 1988). The environmental aspect of ToM is also important, and may reflect the role of language. Ronald and Plomin collected parent, teacher, and self-ratings of the children on three scales: social skills (“Is it important for him to fit in with a peer group?”); communicative skills (“Can she keep a normal two-way conversation going?”); and repetitive behavior (“Does he do the same thing over and over again, in exactly the same way?”). Interestingly, only the children’s communicative skills were significant predictors of children’s performance on false belief tasks. These data reinforce the tight link between linguistic communication and ToM. By contrast, the drive to fit in to a social group may be less dependent on ToM, and may instead reflect social drives that are found across many primate species. Finally, ToM may be relatively independent of “repetitive behavior,” although both measure traits that are atypical in autism. This specific pattern of associations in the Ronald and Plomin study thus fits well with the domain-specific view of ToM arising from the fMRI studies.

Still, the causal mechanism of the observed developmental relationship between language

and ToM remains controversial. According to one view, mental state concepts are intrinsically represented verbally. If complex grammatical structures are necessary for even formulating thoughts about mental states, then ToM development would be parasitic on language ability (the “parasitic” hypothesis). The specific attribute of language commonly implicated in representing another person’s beliefs is the syntax of sentential complements (de Villiers, 2000; de Villiers & de Villiers, 2000; de Villiers & Pyers, 2002). Children’s production and comprehension of this syntactic structure precedes and predicts performance on the false belief task, in both correlation and training studies (de Villiers & de Villiers, 2000; de Villiers & Pyers, 2002; Hale & Tager-Flusberg, 2003).

Alternatively, the way people talk about the mind might be an important source of evidence that children use in constructing mental state concepts. Repeating the same mental state verbs may help draw children’s attention to the abstract similarities between grossly different actions or situations, and between different subjective perspectives. In these ways, verbal communication could facilitate ToM development without playing a constitutive role in representations of mental states (the “communicative” hypothesis; Baldwin & Saylor, 2005; Harris, 2005).

These two hypotheses make distinct predictions for the relationship between language and ToM in adults (and in adult brains). According to the parasitic hypothesis, grammatical constructions provide the representational substrate for thoughts about other minds; the parasitic dependence of ToM on language should be true for adults as well as children. The communicative hypothesis, by contrast, allows for the possibility that language is facilitative during the development of mental state concepts. Once these concepts are formulated, a mature ToM could function even in the absence of language.

The critical test is therefore to investigate the consequences of late acquired aphasia (especially loss of grammatical skills) on an already mature ToM. Two papers in the current issue provide definitive evidence against the parasitic hypothesis. Siegal and Varley were the first authors to report spared ToM competence in two patients with “dense” aphasia (Varley & Siegal, 2000; Varley, Siegal, & Want, 2001). The evidence to date, as they argue in their current review (this issue), is consistent with their proposal that, once mature, ToM does not depend on language.

Aperly and colleagues (this issue) provide strong new evidence for the same hypothesis. They report the performance of PH, a young man who had a left hemisphere stroke, on a battery of language and ToM tests. Although PH is severely impaired on tests of syntax, including specifically the syntax of embedded clauses, he shows no impairments on non-verbal tests of ToM, including second-order inferences (what X thinks that Y thinks). Taken together, these two papers provide clear evidence that adults with a mature ToM can formulate thoughts about other people's thoughts, even in the face of severe grammatical impairments. These data support the view that the link between language and ToM development is primarily communicative, rather than parasitic.

PRECURSORS: INTENTIONAL ACTION AND GAZE DIRECTION

Language is clearly not the only source of information for a developing ToM. A lot of information about others' intentions, perceptions, and feelings is decoded from subtle configurations of human face and body movements (Sabbagh et al., this issue). This information may therefore be processed by key visuo-cognitive "precursors" of ToM: processes for detecting and representing the visual cues to others' intentions and emotions. Central questions about these precursors remain open. For example, are the mechanisms for social perception primarily perceptual or motor in nature? Do they emerge earlier than full-blown mental state attribution? Three papers in the current issue provide new data, and new perspectives, on these key theoretical questions.

One brain region associated with the perception of both intentional action and gaze direction is the right posterior superior temporal sulcus (pSTS; e.g., Allison, Puce, & McCarthy, 2000). In adults, the pSTS has been shown to have a higher BOLD response when watching biological, compared to non-biological motion (Pelphrey et al., 2003). Also, the pSTS distinguishes between motions of human bodies, based on whether the motion is under the control of the actor, or is forced by the physical environment (Morris, Pelphrey, & McCarthy, 2006).

Perhaps most intriguingly, when the visible biological motion is exactly matched, the STS responds differentially based on the relationship between the biological motion and the context of

the action: three studies have found enhanced response in the STS when a hand, body, or eye movement is *unpredicted* given the environmental context (Pelphrey, Singerman, Allison, & McCarthy, 2003; Pelphrey, Morris, & McCarthy, 2004; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004). Senju and colleagues (this issue) provide converging evidence for this signature of perceptual representations of intentional actions. They use stimuli derived from previous fMRI studies (Pelphrey, Viola, & McCarthy, 2004), but identify an ERP component that shows the same functional profile as the pSTS. That is, it is enhanced when an observed character's gaze direction is incongruent with—i.e., unpredicted by—the environment. The authors are then able to identify a similar ERP component in the brains of 9-month-old infants. These results are important in three key respects: (1) this current study provides converging evidence for a perceptual mechanism for processing intentional action that shows an enhanced response for unpredicted actions, while (2) using a new method, EEG, which offers much higher temporal resolution than fMRI, and (3) providing evidence that this mechanism is very early developing, and consistent from infancy to adulthood.

Kaplan and Iacoboni (this issue) investigate another brain region implicated in representations of intentional actions: the right inferior frontal gyrus (IFG). Unlike the pSTS, which is generally thought to form perceptual representations of others' actions, the right IFG is a pre-motor region. The same region is recruited both when subjects *execute* intentional actions, and when—as in the current study—they just *observe* others' intentional actions. The distinction between perceptual and motor representations of others' actions may have intriguing functional consequences. Kaplan and Iacoboni (this issue) report evidence that the IFG representation of actions in context shows the opposite pattern to that commonly observed in the pSTS. For one environmental content (a breakfast table), the response in the IFG was actually higher when the action was *expected* in the context (a precision grip on the handle of a mug, consistent with drinking), compared to an unexpected action (a whole-hand power grip on the body of the mug). Although these results will need confirmation (since the response to precision grips in the IFG is higher across all contexts), the contrast between the functional profiles of the IFG and pSTS may provide important suggestions for the distinct

roles of perceptual and motor representations in understanding others' actions.

The other key claim of Senju and colleagues is that the perceptual mechanisms in the pSTS are already developed in infancy, and are consistent across development. Interestingly, Carter and Pelphrey (this issue) come to a slightly different conclusion about the developmental trajectory of the pSTS. In a pediatric fMRI study, Carter and Pelphrey identify the pSTS by its response preference for biological over non-biological motion. These authors found that while they could identify the pSTS in 7- to 10-year-old children, the selectivity of this region for biological motion developed significantly over this age range. In fact, the youngest children in this study showed no biological motion effect in the pSTS at all. This result will bear replication, given the small number of the youngest children, and the challenges of pediatric neuroimaging. In all, these studies thus point to key new territory for social neuroscience, investigating the developmental trajectory of specialized brain mechanisms for social perception.

COGNITIVE VERSUS AFFECTIVE EMPATHY

Another distinction that plays a key role in many papers is the distinction between cognitive and affective empathy—or between attributing thoughts versus emotions. Affective empathy is composed of both *recognizing* what another person is thinking or feeling (similar to the traditional core of ToM), and *responding* to what another person is feeling with an appropriate emotional state oneself. However, even the recognition of others' emotions might depend on distinct mechanisms from core ToM for two reasons. First, other people's feelings are signaled fairly directly by facial expressions, whereas other people's beliefs and are not expressed in simple behavioral patterns. Second, perceiving someone else experiencing a strong emotion may be more likely to lead to "contagion" than perceiving someone acting on a belief. That is, an intense expression of happiness or fear or sadness might cause an observer to feel a congruent emotion; watching someone else stumble around searching for a missing object is less likely to cause an informed observer to "catch" the other's false belief. Nevertheless, although these considerations make it plausible that ToM and affective

empathy depend on at least partially distinct mechanisms, the neural basis of these mechanisms are only recently being elucidated. Four papers in the current issue contribute to this literature.

The most direct evidence of a distinction between cognitive and affective components of empathy comes from dissociations observed following brain damage to the ventral medial prefrontal cortex (VMPFC). Shamay-Tsoory and colleagues (this issue) developed a novel battery of tasks matched for difficulty, manipulating whether participants were asked to infer a character's belief, or her emotions. Compared to patients with a range of other lesion sites, only the patients with VMPFC lesions showed a selective deficit, scoring lower on tasks that required judgments of emotion. Moreover, the performance of these patients was inversely correlated with independent ratings of the degree of emotional content in individual scenarios, and with the patients' own score on the "Empathic Concern" subsection of the Interpersonal Reactivity Index.

Shamay-Tsoory and colleagues conclude that understanding others' feelings depends on distinct brain mechanisms from understanding others' thoughts, and that this mechanism is located in the VMPFC. Kipps and colleagues (this issue) come to virtually identical conclusions, based on their review of social impairments in frontal-variant fronto-temporal dementia, a progressive neural degenerative disorder that disproportionately affects the ventro-medial frontal cortices. Mojzisch and colleagues (this issue) also converge on the same conclusion. In their current study, they use a range of psychophysiological measures to deepen and complement a previous fMRI study. They report that social-emotional responsiveness in an online interaction is associated with activity in the VMPFC.

By contrast to the above studies, which consider affective empathy as a single coherent unit, Chakrabarti and colleagues (this issue) provide evidence that distinct brain mechanisms respond to distinct emotional expressions. The authors go on to identify brain regions in which there was a correlation between individual differences on the Empathy Quotient (EQ) and neural response to facial expressions of specific emotion. If a single, unified psychological mechanism were responsible for empathy, Chakrabarti and colleagues (this issue) predict that empathy scores and brain activity in response to emotions would converge on the same brain mechanism across emotions. In

fact, Chakrabarti and colleague (this issue) found that distinct brain regions were correlated with EQ during the perception of each specific emotion, although they also found some evidence for a common substrate of empathy across the basic emotions, in the left inferior frontal gyrus. These results provide an interesting contrast with those of Kaplan and Iacoboni (this issue). Both groups tested the correlation between a measure of individual differences in empathy (the EQ scale in Chakrabarti et al.'s study, the Interpersonal Reactivity Index in Kaplan and Iacoboni) and brain activity while observing a human action (facial expressions in Chakrabarti et al.'s study, hand actions in Kaplan and Iacoboni). Both studies identify the IFG as the locus of this correlation. However, Chakrabarti and colleagues observe the correlation in the left IFG (see also Minio-Paluello et al., this issue), while Kaplan and Iacoboni report a right lateralized effect. In all, though, Chakrabarti et al.'s study suggests that the neural substrate of empathy is complex, in that empathy (as measured on the EQ) appears to mediate how different brain regions are used during perception of each specific emotion.

AUTISM

A broad picture is emerging from social neuroscience of the distinct cognitive mechanisms that are necessary for success in social reasoning. A next step may be to use this theoretic landscape in order to better understand conditions affecting social cognition, including autism. Autism, in particular, has historically been associated with deficits in attributing cognitive, rather than affective, states (e.g., Blair 2005). By contrast, Ashwin and colleagues (this issue) report that a sample of adults with autism, relative to matched controls, showed selective deficits in the recognition of negative basic emotions, especially fear. The same participants had much less difficulty recognizing the more "cognitive" expression of surprise. The authors conclude that autism may in part reflect abnormalities in amygdala function. Mazzola and colleagues (this issue) make a parallel claim for Turner's syndrome. Establishing what is specific to the autistic brain is outside the scope of these two papers.

Equally important for characterizing the social impairments in autism, though, is the program of identifying components of social reasoning that are not impaired. Leslie and colleagues (this

issue) report that while children with autism fail false belief tasks, they nevertheless pass simple tests of moral reasoning. In particular, the children with autism were just as good as typically developing 4-year-olds at identifying that it is bad to steal someone else's cookie when it makes them cry, but it is not bad to eat one's own cookie, even if the other person (who greedily wants two cookies) starts to cry. This finding bears out the frequent clinical reports that even among those with high-functioning autism or Asperger syndrome, while they may have deficits on advanced ToM tests (Baron-Cohen, 1997; Baron-Cohen et al., 2000), they may care passionately about moral issues such as social justice. In this sense, their empathy deficits do not produce an outcome of psychopathy. They are rarely knowingly cruel or hurtful and, even if they have difficulties in reading the other's mind or facial expressions of emotion, they have a strong sense of right and wrong. The neuroscience of moral development may turn out to be a key way of differentiating the two empathy conditions of autism and psychopathy.

CONCLUSIONS

As editors, we are struck by the range of methods that have just recently been brought to bear on the psychological and neural basis of ToM. In this single issue, we have gathered papers that use TMS, EEG, psychophysiology, eye-tracking, and fMRI, and that investigate ToM in typically developing children and children with autism, in twins, in women with Turner's syndrome, in patients with acute lesions, and in patients with neurodegenerative disorders.

Nevertheless we are aware that the neuroscience of ToM is far broader and deeper than can be covered by the papers we have assembled in this Special Issue. For example, if ToM is partly genetic, we will need molecular genetic studies to identify which candidate genes may be functionally related to the development of ToM. This will be an important area for future research. Second, there is clear reason to believe that the development of ToM and empathy may receive an important contribution from androgens as well as genes (Baron-Cohen, 2003; Mazzola et al., this issue). Since the brain regions implicated in ToM are all rich in androgen receptors, this may prove to be a rich source of variation in individual differences in ToM development. Finally, the role

of the early environment in supporting or damaging the development of empathy, as was first proposed by Bowlby (1969), will be an important area of study towards producing an integrated model of the development of ToM. We hope this Special Issue not only highlights the progress that is being made in the understanding of ToM, but also the existing gaps in this field that a new generation of social neuroscientists will fill.

REFERENCES

- Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *Neuroimage*, 30(3), 1059–1068.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Science*, 4(7), 267–278.
- Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W. L., & Humphreys, G. W. (2006). Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence for consistent performance on non-verbal, “reality-unknown” false belief and false photograph tasks. *Cognition* [Epub ahead of print].
- Astington, J. W., & Baird, J. A. (Eds.). (2005). *Why language matters for theory of mind*. Oxford, UK: Oxford University Press.
- Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory of mind development. *Developmental Psychology* 35(5), 1311–1320.
- Baldwin, D. A., & Saylor, M. M. (2005). Language promotes structural alignment in the acquisition of mentalistic concepts. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 123–143). Oxford, UK: Oxford University Press.
- Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Baron-Cohen, S. (2003). *The essential difference: Male and female brains and the truth about autism*. Cambridge, MA: Basic Books.
- Baron-Cohen, S., Baldwin, D., & Crowson, M. (1997). Do children with autism use the speaker’s direction of gaze strategy to crack the code of language? *Child Development*, 68(1), 48–57.
- Baron-Cohen, S., Leslie, A. M., & Frith, H. (1985). Does the autistic child have a theory of mind? *Cognition*, 21, 37–46.
- Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (Eds.). (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience*. Oxford, UK: Oxford University Press.
- Blair, R. J. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, 14(4), 698–718.
- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77(1), B25–31.
- Bowlby, J. (1969). *Attachment and loss* (Vol. 1). London: Hogarth Press.
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children’s theory of mind. *Child Development*, 72(4), 1032–1053.
- Charman, T., & Baron-Cohen, S. (1992). Understanding drawings and beliefs: A further test of the metarepresentation theory of autism: A research note. *Journal of Child Psychology and Psychiatry*, 33(6), 1105–1112.
- Cosmides, L., & Tooby, J. (1997). Dissecting the computational architecture of social inference mechanisms. *Ciba Foundation Symposium*, 208, 132–156.
- de Villiers, J. (2000). Language and theory of mind: What are the developmental relationships? In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience*. Oxford, UK: Oxford University Press.
- de Villiers, J., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. J. Riggs (Eds.), *Children’s reasoning and the mind* (pp. 191–228). Hove, UK: Psychology Press.
- de Villiers, P. A. (2005). The role of language in theory-of-mind development: What deaf children tell us. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 266–297). Oxford, UK: Oxford University Press.
- de Villiers, J., & 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.
- Dennett, D. (1978). Beliefs about beliefs. *Behavioral and Brain Science*, 1, 568–570.
- Dunn, J., & Brophy, M. (2005). Communication, relationships, and individual differences in children’s understanding of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 50–69). Oxford, UK: Oxford University Press.
- Figueras-Costa, B., & Harris, P. L. (2001). Theory of mind development in deaf children: A nonverbal test of false-belief understanding. *Journal of Deaf Studies and Deaf Education*, 6(2), 92–102.
- Folstein, S. E., & Rutter, M. L. (1988). Autism: Familial aggregation and genetic implications. *Journal of Autism and Developmental Disorders*, 18(1), 3–30.
- Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science*, 6(3), 346–359.
- Harris, P. L. (2005). Conversation, pretense, and theory of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 70–83). Oxford, UK: Oxford University Press.

- Leslie, A., & Thaiss, L. (1992). Domain specificity in conceptual development. *Cognition*, *43*, 225–251.
- Moeller, M. P., & Schick, B. (2006). Relations between maternal input and theory of mind understanding in deaf children. *Child Development*, *77*(3), 751–766.
- Morgan, G., & Kegl, J. (2006). Nicaraguan Sign Language and theory of mind: The issue of critical periods and abilities. *Journal of Child Psychology and Psychiatry*, *47*(8), 811–819.
- Morris, J. P., Pelphrey, K. A., & McCarthy, G. (2006). *Perceived agency influences brain activity evoked by biological motion*. Manuscript submitted for publication.
- Pelphrey, K. A., Mitchell, T. V., McKeown, M. J., Goldstein, J., Allison, T., & McCarthy, G. (2003). Brain activity evoked by the perception of human walking: Controlling for meaningful coherent motion. *Journal of Neuroscience*, *23*(17), 6819–6825.
- Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, *16*(10), 1706–1716.
- Pelphrey, K. A., Singerman, J. D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: The influence of context. *Neuropsychologia*, *41*(2), 156–170.
- Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science*, *15*(9), 598–603.
- Peterson, C. C., & Siegal, M. (1999). Deafness, conversation and theory of mind. *Journal of Child Psychology and Psychiatry*, *36*(3), 459–474.
- Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and US pre-schoolers. *Psychological Science*, *17*(1), 74–81.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*(2), 235–239.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporoparietal junction in “theory of mind”. *NeuroImage*, *19*(4), 1835–1842.
- Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, *42*(11), 1435–1446.
- Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and “theory of mind” in an agrammatic aphasic patient. *Current Biology*, *10*(12), 723–726.
- Varley, R., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, *7*(6), 489–493.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.
- 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.
- Zaitchik, D. (1990). When representations conflict with reality: The preschooler’s problem with false beliefs and “false” photographs. *Cognition*, *35*(1), 41–68.