

What's domain-specific about theory of mind?

Valerie E. Stone

University of Queensland, St Lucia, Australia

Philip Gerrans

University of Adelaide, Adelaide, Australia

Twenty years ago, Baron-Cohen and colleagues argued that autistic performance on false belief tests was explained by a deficit in metarepresentation. Subsequent research moved from the view that the mind has a domain-general capacity for metarepresentation to the view that the mind has a domain-specific mechanism for metarepresentation of mental states per se, i.e., the theory of mind mechanism (ToMM). We argue that 20 years of data collection in lesion patients and children with autism supports a more parsimonious view closer to that of the 1985 paper. Lower-level domain-specific mechanisms—e.g., tracking gaze, joint attention—interacting with higher-level domain-general mechanisms for metarepresentation, recursion, and executive function can account for observed patterns of deficits in both autism and neurological patients. The performance of children with autism or orbitofrontal patients on ToM tests can be explained more parsimoniously by their deficits in lower-level domain-specific mechanisms for processing social information. Without proper inputs, the intact capacity for metarepresentation by itself cannot make correct ToM inferences. Children with autism have no impairment in false photograph tests because their metarepresentational capacity is intact and they have no impairment in inputs required for such tests. TPJ patients have equivalent deficits on ToM and non-ToM metarepresentational tasks, consistent with a failure in domain-general processing. If deficits on ToM tasks can result from deficits in low-level input systems *or* in higher-level domain-general capacities, postulating a separate ToM mechanism may have been an unnecessary theoretical move.

In 1985, Simon Baron-Cohen, Alan Leslie, and Uta Frith published an influential paper entitled “Does the autistic child have a theory of mind?” They argued that a child who lacked a capacity for metarepresentation would be unable to apply it to the developmentally vital task of understanding the representational nature of other's thoughts, and that the abnormal developmental trajectory of autism would result. They presented evidence that children with autism were more likely to fail “false belief” tests, which require the metarepresentation of mental states, than were typically developing children. These findings have been replicated and extended over the past two decades (Wellman, Cross, & Watson, 2001). Work

on autism in general has expanded greatly over the same time, with many accounts of the core deficit in autism emerging as alternatives to the theory of mind account: deficits in executive function, emotion and empathy, central coherence, attentional effects following from cerebellar abnormalities (e.g., Baron-Cohen, 2004; Baron-Cohen & Wheelwright, 2004; Courchesne et al., 1994; Happé, 1999; Ozonoff, Pennington, & Rogers, 1991; Schultz, 2005). As a result, the empirical database on early deficits in autism is much richer than it was two decades ago, and we believe it is time to revisit the question of domain-specificity in theory of mind in light of new evidence.

Correspondence should be addressed to: Valerie E. Stone, School of Psychology, McElwain Bldg., University of Queensland, St Lucia, QLD 4072, Australia. E-mail: stone@psy.uq.edu.au

The original paper by Baron-Cohen et al. says specifically that theory of mind (ToM) is “one of the manifestations of a basic metarepresentational capacity” (1985, p. 37, emphasis added). “Metarepresentation” is used in the literature to mean representing the relation between representation and referent. Following the spirit of that argument, several theorists have put forth a case for metarepresentation being a domain-general capacity that includes but is not limited to metarepresentation of mental states (Corballis, 2003; Perner, 1991; Stone & Gerrans, 2006; Suddendorf, 1999; Suddendorf & Whiten, 2001). In contrast, other developmental theorists have pursued the idea that autism is a failure of a *domain-specific* capacity for the metarepresentation of mental states underwritten by a specific cognitive mechanism: the Theory of Mind Mechanism (ToMM; Leslie, 1987, 1994). As children with autism often have difficulty with metarepresentation of mental states, but not other metarepresentational tasks (Leslie & Thaiss, 1992), it seemed necessary to propose such a specific module. Evolutionary psychologists embraced the idea of such a module with enthusiasm, as it represented a high-level domain-specific mechanism with a specific adaptive function, the metarepresentation of mental states (Cosmides & Tooby, 1995). Given that neuroscience is one important source of evidence for modularity, neuroscientists began the search for evidence of specialized brain regions that instantiate this module, and for evidence of the dissociation from other cognitive functions that would be the hallmark of modularity (e.g., Saxe, Carey, & Kanwisher, 2004; Stone, Baron-Cohen, & Knight, 1998). Those researchers who dissented from the view that children with autism have a metarepresentation deficit that is specific to mental states are the very same theorists who did not espouse the modular view (Perner, 1991; Russell, Saltmarsh, & Hill, 1999).

The concept of “modularity” has been interpreted in several different ways. On some views modularity does not imply anatomical localization. A distributed circuit can be specialized for performance of a particular cognitive function (Atkinson & Wheeler, 2004; Coltheart, 1999; Stone et al., 1998). On any view of domain specificity, however, functional specialization must depend on specialized neural circuitry, distributed or localized. Both fMRI and lesion studies can show that particular sets of neurons are necessary for performance of a certain

cognitive function. However, from the fact that a set of neurons is *necessary* for performance of a particular cognitive function it does not follow that the neural circuit is *specific* to that function. It may be necessary for other cognitive functions as well. Working memory, which has a particular neural instantiation, is necessary for understanding complex grammatical sentences; however, working memory and its neural substrate are not specialized for that function. The boundaries of domain-specific mechanisms (functional and anatomical) are drawn on the basis of both necessity *and* specialization for particular tasks. This minimal characterization of domain-specificity makes no claims (as modular theorists do) about speed, automaticity, encapsulation, localization, and representational format, which do not concern us here. When we discuss a domain-specific mechanism, we mean a neural circuit that is not necessarily localized but is both necessary for, and specific to, a particular cognitive task.

We argue that the postulation of a domain-specific ToM mechanism represents an unnecessary move in theoretical understanding of both autism and normal development. An alternative model of ToM can account for the data from both autism and neurological patients more parsimoniously. This alternative view, however, will need to account for the data which led to the postulation of a ToM mechanism in the first place: the apparent dissociation between metarepresentation *per se* and metarepresentation of mental states.

Let us begin by stating what is in common between all theorists of ToM development. First, development of ToM depends on the *prior* development of a suite of specialized lower-level cognitive mechanisms. These precursor mechanisms represent vital information about the social world of the infant and toddler and mediate her/his earliest interactions with others. These mechanisms enable face processing, emotion processing, representations of gaze direction, gaze monitoring, detection of animacy, tracking of intentions and goals, and joint attention (Baron-Cohen, 1995, 2004; Charman, Baron-Cohen, Swettenham, Baird, Cox, & Drew, 2000; Crichton & Lange-Küttner, 1999; Csibra, Biro, Koos, & Gergely, 2003; Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998; Saxe et al., 2004; Schultz, 2005; Stone, 2005, 2006; Wellman, Phillips, Dunphy-Lelii, & LaLonde, 2004; Woodward, 1999). These capacities appear to be domain specific. They seem to be specific to social stimuli, to be

shared with other primates, and to depend on neural circuitry that responds preferentially to social stimuli (e.g., Blakemore, Boyer, Pachot-Clouchard, Meltzoff, Segebarth, & Decety, 2003; Campbell, Heywood, Cowey, Regard, & Landis, 1990; Hare, Call, Agnetta, & Tomasello, 2000; Kumashiro, Ishibashi, Itakura, & Iriki, 2002; Perrett et al., 1990). Gaze monitoring, for example, seems to involve specific regions of the superior temporal sulcus that respond to the stimulus of eye-gaze direction but not to other non-social stimuli, even other stimuli that are physically similar to images of eyes (Campbell et al., 1990; Hoffman & Haxby, 2000). Similar preferential neural responses are involved in assessing others' goals and intentions, which seems to depend on specific representations of certain movement patterns: limb movement combined with gaze, head, or body orientation (Blakemore et al., 2003; Jellema, Baker, Wicker, & Perrett, 2000).

The development of these capacities ensures that the normal toddler is equipped with a sophisticated battery of domain-specific mechanisms that enable her/him to negotiate the social world on the basis of perceptually available information. In addition, the toddler appears to have a domain-general capacity for secondary representation (the capacity to hold in mind and compare two different representations), but cannot metarepresent beliefs (Suddendorf, 1999; Suddendorf & Whiten, 2001). In this respect she/he has similar social capacities to the great apes, using essentially the same cognitive equipment (Suddendorf & Whiten, 2001, 2003).

Leslie (1987, 1992) has maintained that pretense in toddlers is evidence of early mental state metarepresentational abilities, and one recent study credits 15-month-olds with this ability (Onishi & Baillargeon, 2005). Other developmental research questions these claims, arguing that they can be explained by secondary representations or understanding pretense as a special category of action (cf. Sobel & Lillard, 2002; Suddendorf & Whiten, 2001; Wellman & Lagattuta, 2000). Developmental research does not address the question of domain-specificity directly, unless it can be shown that: (1) a ToM task definitely depends on metarepresentation, which is debated in the case of pretense; and (2) infants can solve a ToM metarepresentational task before they can solve any other well-matched metarepresentational task, which has not been established in the case of 15-month-olds. Thus,

just as debates about *where* ToM is in the brain do not resolve the question of domain specificity, neither do current debates about *when* ToM metarepresentation emerges.

Another point of agreement among ToM theorists is that by age four, the typically developing child is equipped with a battery of higher level *domain-general* cognitive mechanisms that serve metacognitive computational functions. Advanced executive function (meaning working memory, inhibition and flexible control of attention), secondary representations, recursion (the ability to compute embedded representations), and metarepresentation (understanding the representational nature of the relationship between representation and referent) all come on line by age four, though they may develop further beyond that age (De Villiers & Pyers, 2002; Perner, 1991; Smith, Apperly, & White, 2003; Suddendorf, 1999; Suddendorf & Whiten, 2001). A further point agreed on by many theorists is that by age four a child has some level of semantic knowledge about what mental states are like (Gopnik & Wellman, 1992; Leslie, 1987; Wellman, 1988). This semantic knowledge includes an understanding that beliefs and desires are private and changeable independent of the external state of reality changing (Wellman & Lagattuta, 2000). Positing semantic knowledge of a particular content domain, however, is quite different from positing that there are specialized computational processes for a certain domain. Semantic knowledge of a particular content domain is not a "domain-specific mechanism."

The original paper by Baron-Cohen et al. (1985) suggested that passing the false belief test engaged some or all of these metacognitive mechanisms. The ToMM theory suggests that *as well as these domain-general mechanisms* the child develops a domain-specific mechanism for the metarepresentation of mental states. See Figure 1 for a graphic representation. A child equipped with both a ToMM and domain-general mechanisms essentially has two metarepresentational devices: one for social metarepresentations and one for metarepresentations in other domains. We think that it is unlikely that two separate metarepresentational mechanisms (a) evolved twice and (b) develop twice.

One of the main challenges to the domain-specific view of ToM deficits in autism has been the view that deficits in a domain-general ability, executive function (EF), can account for autistic children's failures on ToM tasks (e.g., Ozonoff

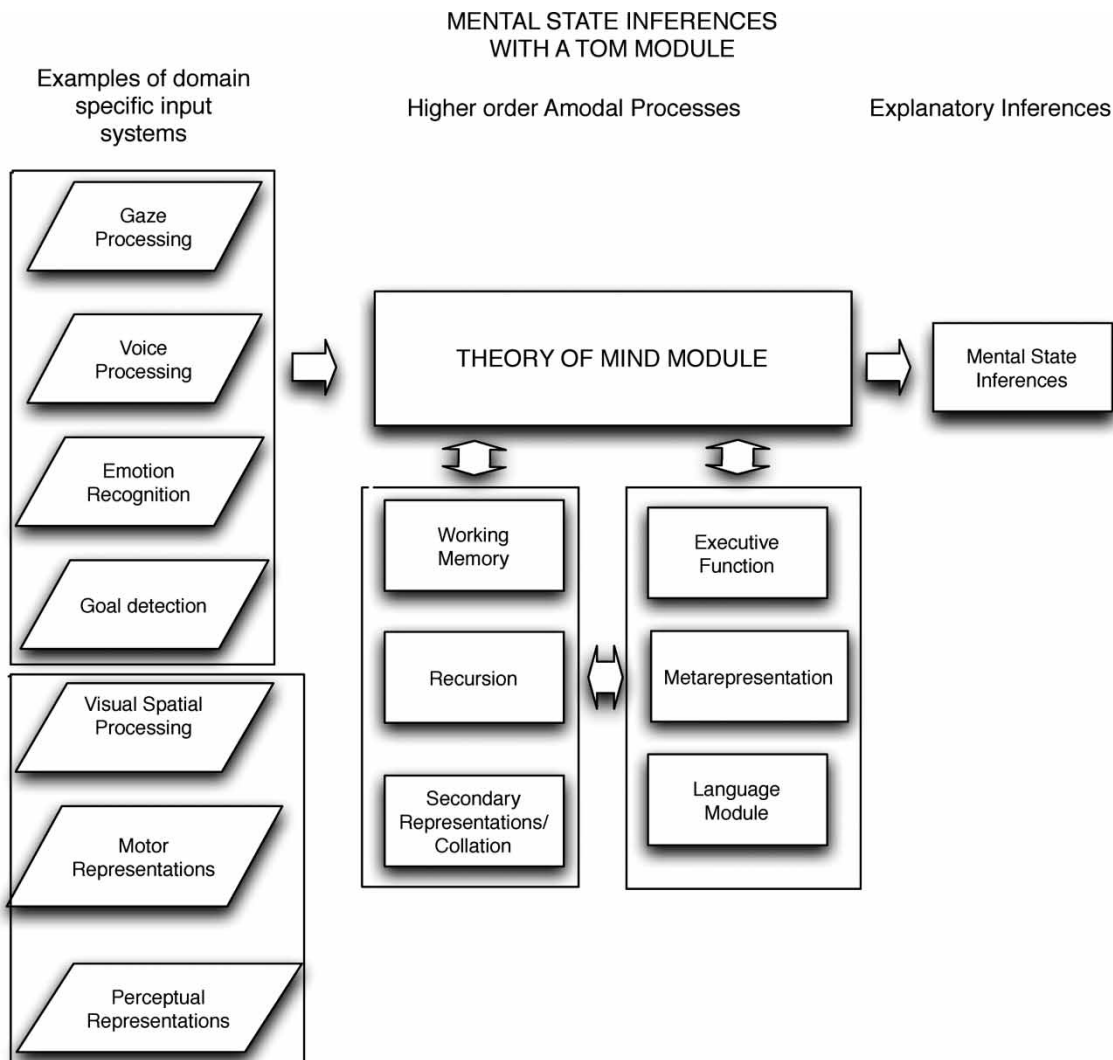


Figure 1. Mental state inferences with a ToM module. Architecture postulated by the ToMM theory. A child equipped with both a ToMM and domain-general mechanisms essentially has two metarepresentational devices: one for social metarepresentations and one for metarepresentations in other domains.

et al., 1991). We wish to claim something quite different, namely that ToM abilities depend, not only on EF, language, recursion, or metarepresentation per se, but on their developmental and on-line interaction with the low level precursor mechanisms previously described, e.g., gaze processing, emotion recognition. Failure of these low-level abilities can also cause failure on ToM tasks, even in subjects with intact EF. This view might explain why recent empirical results show that early deficits in EF in autism cannot always account for the difficulties children with autism have with ToM tasks. Toddlers with autism show evidence of joint attention deficits, but not always early EF deficits (Griffith, Pennington, Wehner,

& Rogers, 1999; Rutherford & Rogers, 2003). Furthermore, deficits in EF are not always apparent when children with autism are tested by a computer rather than a person (Ozonoff, 1995). In our view, this pattern is to be expected if the cause of the failure is a deficit of low-level input systems.¹ How, then, can one account for the fact that children with autism show deficits in the metarepresentation of mental states, but not in other metarepresentational tasks?

¹ Indeed, much recent work in explaining autism has looked at how early social deficits contribute to later theory of mind difficulties (e.g., Baird et al., 2000; Baron-Cohen, 1995; Boucher & Lewis, 1992; Dawson et al., 2004a, 2004b; Rutherford & Rogers, 2003; Schultz, 2005).

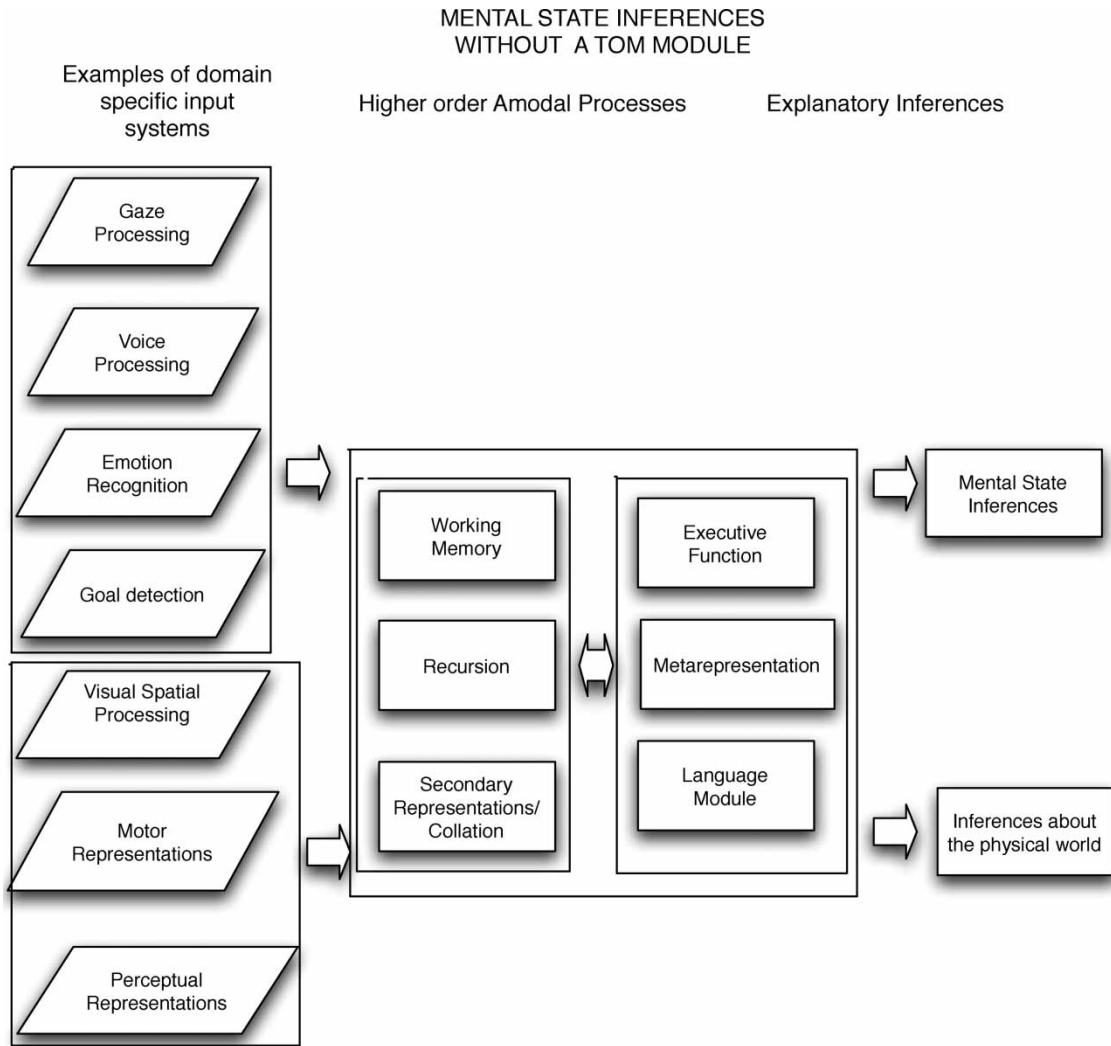


Figure 2. Mental state inferences without a ToM module. ToM is simply an interaction between low-level precursor domain-specific mechanisms and high-level domain-general mechanisms. The result of this interaction is not the creation of an extra domain-specific mechanism, as in Figure 1, but the wiring up of distributed metarepresentational circuitry that can take social information as input and deliver ToM inferences as output. The performance of children with autism on ToM tests can be explained more parsimoniously by the view that they have deficits, not in metarepresentation, but in lower-level domain-specific mechanisms for processing social information. Without proper inputs, the intact capacity for metarepresentation by itself cannot make correct ToM inferences.

The interaction between low-level precursor domain-specific mechanisms and high-level domain-general mechanisms can account for normal performance on ToM tasks (Gerrans, 2003; Stone & Gerrans, 2006). The result of this interaction is not the creation of an extra domain-specific mechanism, but the wiring up of distributed metarepresentational circuitry that can take social information as input and deliver ToM inferences as output. See Figure 2 for a graphic representation of this possible architecture. Metarepresentation operating on information about

eye gaze² and attention (who saw or was attending to what) allows us to represent others' knowledge states (who knew what) (Baron-Cohen, 1995). Without the necessary inputs, no proper inferences are made. Recursion operating on metarepresentations of mental states can allow us to reason about not just others' thoughts, but

² To say that eye gaze is *an* input to ToM inferences is not to say that it is the only necessary input. Obviously, other senses can compensate, as they do in enabling blind children to make ToM inferences, even if their development is slightly delayed (Baron-Cohen, 1995).

also others' thoughts about thoughts (Corballis, 2003). Indeed, children's ability to use embedded syntactical structures comes on line shortly *before* the ability to solve false belief tasks, supporting the view that recursion may be one general ability serving both tasks (De Villiers & Pyers, 2002; Smith et al., 2003). Executive function allows one to keep the elements of a social interaction in mind, and inhibit one's own knowledge of the state of reality when asked what someone else's mental state is (Stone, 2005). Both working memory and inhibition have been shown to be related to performance on ToM tasks (Carlson & Moses, 2001; Keenan, 1998; Stone et al., 1998). Thus, deficits on ToM tasks can result from deficits in *either* low-level input systems (e.g., joint attention) *or* higher-level domain-general capacities rather than in a separate ToM module.

It is worth pointing out that recursive embedding, e.g., of a clause within a sentence, is not the same thing as metarepresentation, which is the representation of a representational relationship. Metarepresentation requires the recursive embedding of representational relationships. Recursion provides a schema but does not explain how elements embedded within that schema are generated. That metarepresentation and recursion are separate domain-general abilities explains why chimps who understand recursive dominance relationships still cannot metarepresent mental states: they do not understand representational relationships (Suddendorf & Whiten, 2001, 2003). Syntactic recursion seems closely tied to ToM in development (De Villiers & Pyers, 2002; Smith et al., 2003) but this does not mean either that the two abilities are identical or that ToM is domain specific.

The performance of children with autism on ToM tests can be explained more parsimoniously by the view that they have deficits, not in metarepresentation, but in lower-level domain-specific mechanisms for processing social information. Without proper inputs, the intact capacity for metarepresentation by itself cannot make correct ToM inferences. Autistic children have no impairment on false photograph tests because they have no impairment in inputs required for those tests. See Figure 3 for a graphic representation. Thus, we argue that autistic deficits on ToM tasks are evidence *not* of impaired metarepresentation, but are instead evidence of impaired *inputs* to ToM inferences. Research on social deficits in autism has shown clearly that children with autism have deficits in many early domain-

specific social competences: face recognition, facial expression recognition, processing of gaze direction, and joint attention (Baird, Charman, Cox, Swettenham, Wheelwright, & Drew, 2000; Baron-Cohen, 1995; Boucher & Lewis, 1992; Dawson et al., 2004a; Dawson, Webb, Carver, Panagiotides, & McPartland, 2004b; Rutherford & Rogers, 2003; Schultz, 2005). These deficits in lower-level domain-specific mechanisms have been proposed as causes of the ToM impairment seen in autism, because the causal antecedents of ToM have not developed properly. Strong evidence against our view would be a case of an individual with deficits on metarepresentational ToM tasks, but no lower-level social deficits. We note that no such case has yet been demonstrated in the autism literature.

We suggest that metarepresentation in autism is intact and that evidence supports this view. In the absence of comorbid intellectual disability, individuals with autism seem to have intact capacities for both metarepresentation and recursion. They perform well on a non-mental metarepresentational task, the false photograph test, and can sometimes pass false belief tests (Baron-Cohen, 1989; Baron-Cohen, Wheelwright, Stone, Jones, & Plaisted, 1999a; Leslie & Thaiss, 1992). Baron-Cohen, Wheelwright, Stone, and Rutherford (1999b) report three cases of high-functioning individuals with ASD, a mathematician, a physicist and an engineering student at a prestigious university. All three fields require expertise in mathematics, that most recursive of cognitive abilities. Mathematics requires people to represent the relations between symbols and their objects (magnitudes, spatial relations) and to perform recursive computations over these symbols. These three men excelled in their fields (in fact, the mathematician was a Fields medalist), showing that their capacities for metarepresentation, recursion, and EF were intact. All did well on the Tower of Hanoi, an EF test. However, all three had difficulty in inferring what someone was feeling or paying attention to from pictures of the eye region of the face (Baron-Cohen et al., 1999b), indicating a problem with lower-level domain-specific mechanisms for face and gaze processing rather than with metarepresentation. There is currently no evidence for a *pure* metarepresentational ToM deficit in autism.

In our view, the two routes to deficits on ToM tasks are deficits in low-level input systems (e.g., representations of gaze, joint attention) *or* in higher-level domain-general capacities (executive

FALSE PHOTO INFERENCES

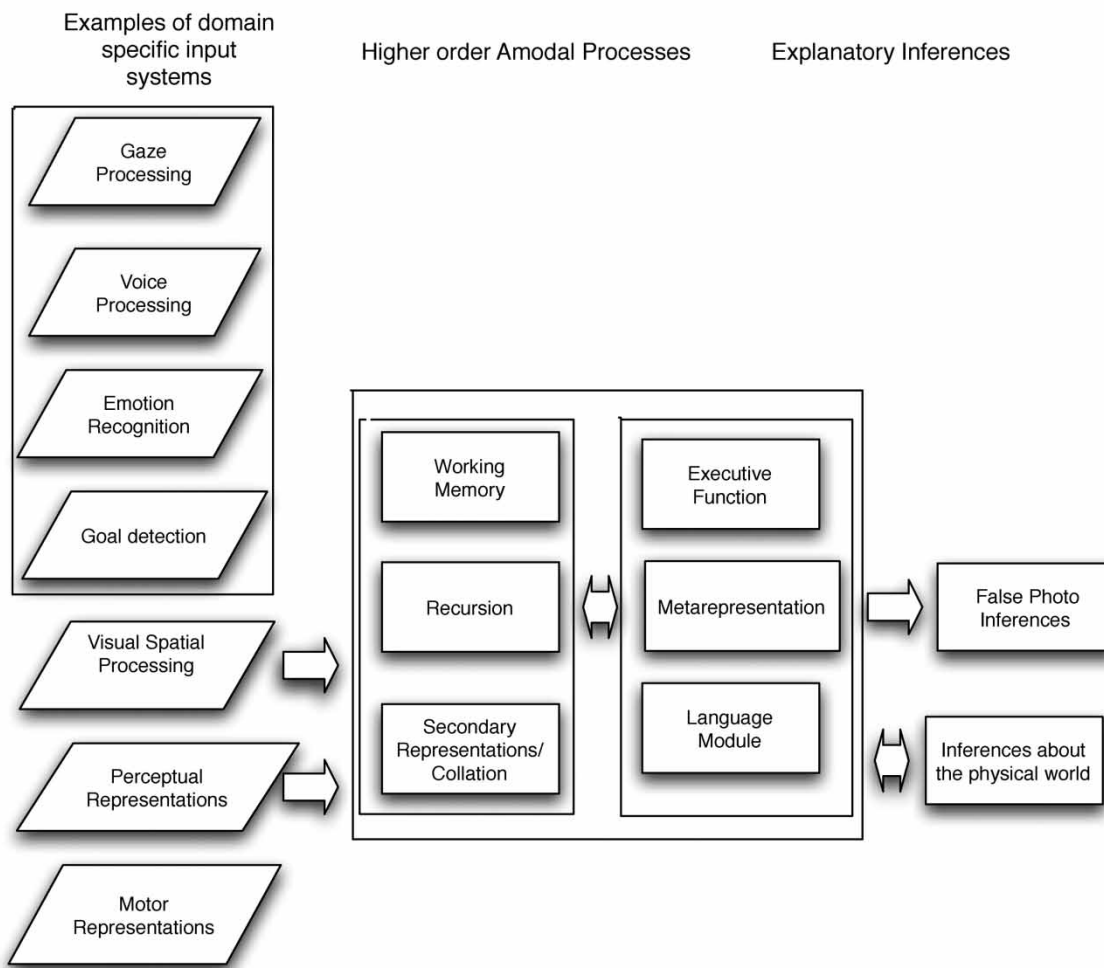


Figure 3. Architecture supporting false photograph inferences. Children with autism have no impairment on false photograph tests because their metarepresentational capacity is intact and they have no impairment in inputs required for such tests.

function, metarepresentation, and recursion). In the spirit of empiricism, we offer to be proven wrong. An individual with a deficit in ToM metarepresentation without any deficit in lower-level social competences and without any more general deficit in other types of metarepresentation, executive function, or recursion would provide conclusive evidence against our view. No such case from autism has yet been demonstrated in the literature.

If children with autism do not provide such evidence, then perhaps data from neurological patients with lesions and selective deficits would provide stronger evidence. Existing evidence from neuropsychology, however, seems only to support the view that it is not possible to be impaired on ToM tasks without a concurrent deficit in either lower-level social abilities or

other domain-general abilities. Patients with orbitofrontal cortex (OFC) damage, for example, have been found to be impaired on ToM tasks (Gregory et al., 2002; Stone et al., 1998). However, they are also impaired at recognizing facial expressions and at judging mental states from eye gaze or expression in the eye region of the face (Gregory et al., 2002; Hornak, Rolls, & Wade, 1996; Snowden et al., 2003). Furthermore, even their problems with some ToM tasks may result from a difficulty in tracking intentions rather than beliefs, as these are ToM tasks that, unlike false belief tasks, do not specifically require metarepresentation of belief (Stone, 2005). Thus, their mentalizing deficits can be explained as a result of deficits in lower-level domain-specific mechanisms, not in higher-order domain-general mechanisms. Indeed, some can perform at ceiling on false

belief tasks, even those requiring 2nd- and 3rd-order mental state inferences (Stone, 2005; Stone et al., 1998). Patients with medial frontal damage who have ToM deficits all have accompanying executive function deficits, and extensive medial frontal damage does not necessarily cause impairment in ToM (Apperly, Samson, Chiavarino, & Humphreys, 2004; Bird, Castelli, Malik, Frith, & Husain, 2004; Gregory et al., 2002; Happé, Mahli, & Checkley, 2001; Stone, 2005).

There has been much recent excitement over the possibility that patients with temporoparietal junction (TPJ) lesions might represent a group with specific ToM deficits (Apperly et al., 2004; Samson, Apperly, Chiavarino, & Humphreys, 2004). However, the most recent evidence from such patients indicates that these deficits are not ToM specific either (Apperly, Samson, Chiavarino, Bickerton, & Humphreys, *in press*). Language and EF deficits can account for deficits on language-based false belief tasks that require inhibiting knowledge of the correct state of reality, although some patients with lesions in the same areas did well on these same tasks (Apperly et al., 2004; Samson et al., 2004). More interesting is the finding that other TPJ patients who failed to perform above chance on nonverbal false belief tasks also failed to perform above chance on a very closely matched nonverbal test of non-mental metarepresentation, the false photograph test (Apperly, Samson, & Humphreys, 2005; Apperly et al., *in press*). Thus, there is as yet no evidence from neurological patients that supports the claim that a capacity for ToM inferences can be impaired independently of damage to lower-level input mechanisms or higher-level domain-general mechanisms.

However, we do not believe that excitement over the role of TPJ was unwarranted. Rather, the TPJ might be interesting for a different reason: Apperly et al.'s (*in press*) data suggests that metarepresentation might be a crucial ability subserved by TPJ. These patients also have deficits in language, another ability that uses metarepresentation (Stone, 2006). Although TPJ does not appear to be the locus of the fabled ToM module, it may be the locus of a much more interesting ability. Metarepresentation, some researchers have argued, is one of the key abilities that separates humans from the great apes, one of the key cognitive capacities that emerged over the course of hominid evolution (Corballis, 2003; Suddendorf, 1999). Thus, excitement over the importance of this region should be excitement

over looking at a region subserving a domain-general ability rather than a domain-specific ability.

Some researchers, however, have argued that neuroimaging data gives support to the domain-specific view. Recent results indicate that the TPJ is differentially active for false belief tasks and false photograph tasks. This apparent discrepancy would be resolvable with some additional analyses of current imaging data. Saxe and Kanwisher (2003) present contrasts between false belief and false photograph trials, and between false belief task and non-metarepresentational task trials. They do not, however, present any contrasts for the whole brain between false photograph task and non-metarepresentational task trials.³ If such a contrast did show differential activation in TPJ, then this would support the view that TPJ might underwrite metarepresentation more generally. In that case, the differential activation in TPJ between false belief and false photograph trials might simply reflect how much each task demands of a metarepresentational capacity, though both require metarepresentation. In contrast, if TPJ was not differentially active when non-metarepresentational tasks are subtracted from false photograph tasks, there would be stronger evidence for TPJ's subserving belief state metarepresentation specifically.

A variety of tasks have been used to measure ToM in neuroscience (Stone, 2005). False belief tasks are considered the strongest tests because they require metarepresentation. There are some differences, however, in the false belief tasks used by different experimenters. Saxe and Kanwisher (2003) used verbal false belief tasks and verbal false photograph tasks, tasks that differed in difficulty for their participants. Apperly et al. (2005, *in press*) have used nonverbal false belief and false photograph tasks closely equated for

³ Figure 4 of Saxe and Kanwisher (2003) compares how certain voxels respond to particular types of stimulus stories: false belief, false photograph, physical descriptions of people, desire, and nonhuman. However, these voxels were preselected to be more active in false belief than in false photograph tasks in "more than half" of the 14 participants. Thus, not all voxels were compared on all tasks. Furthermore, for some subsets of these 14 participants, there was no significant difference in the activation between false belief and false photograph tasks, a situation that is, on some interpretations, at odds with ToM being a universal domain-specific mechanism. The comparison we are suggesting (false photograph vs. non-metarepresentational tasks) needs to be done for all participants for all voxels.

difficulty in neurotypical participants. Convergence between neuroimaging and lesion studies depends on researchers using the same paradigms. These results will be clearer when a uniform paradigm, verbal or nonverbal, is used in both types of research.

In the absence of strong evidence for dedicated neural circuitry or the dissociation of ToM from metarepresentational or lower-level deficits, it seems that the interaction of several domain-general mechanisms and lower-level domain-specific mechanisms can account for the flexibility and sophistication of behavior, which has been taken to be evidence for a domain-specific ToM mechanism. One can understand, as an episode in the history of science, why it seemed necessary to posit a specific ToM mechanism. Early evidence from autism seemed to show a pattern of impairment in metarepresentation of belief, without a corresponding impairment in metarepresentation in other domains (Baron-Cohen et al., 1985; Frith, Morton & Leslie, 1991; Leslie & Thaiss, 1992). Only later did it become clear that the primary deficits in autism were at a lower level. Furthermore, early explorations of the idea of metarepresentation primarily centered around its role in the embedding of beliefs (Dennett, 1987; Leslie, 1987). The relevance of metarepresentation and recursion for other cognitive functions—language, mathematics, episodic memory, and future planning (Corballis, 2003; Perner, 1991; Suddendorf, 1999)—was obscured by the significance of the apparent dissociation between ToM and domain-general metarepresentation.

The history of science is replete with examples of theorists postulating unnecessary entities. In insisting that the earth was at the center of the universe, Ptolemy proposed the notion of epicycles to explain the movements of the planets across the heavens: separate circular orbits in which the planets moved whose center was attached to the celestial spheres, which surrounded the earth. Copernicus and Kepler, moving the sun to the center, could dispense with this unnecessary theoretical addition. The positing of a separate ToM module appears to have been an “epicycle,” an unnecessary addition to psychological theorizing. We argue that metarepresentation, as a *domain-general ability*, should be moved back to its rightful place at the center of our species’ uniquely human cognitive abilities.

REFERENCES

- Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W., & Humphreys, G. W. (in press). 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*.
- 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*(10), 1773–1784.
- Apperly, I. A., Samson, D., & Humphreys, G. W. (2005). Domain-specificity and theory of mind: Evaluating neuropsychological evidence. *Trends in Cognitive Sciences*, *9*(12), 572–577.
- Atkinson, A. P., & Wheeler, M. (2004). The grain of domains: The evolutionary-psychological case against domain-general cognition. *Mind & Language*, *19*(2), 147–176.
- Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., & Drew, A. (2000). A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(6), 694–702.
- Baron-Cohen, S. (1989). The autistic child’s theory of mind: A case of specific developmental delay. *Journal of Child Psychology & Psychiatry*, *30*(2), 285–297.
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Baron-Cohen, S. (2004). The cognitive neuroscience of autism. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*(7), 945–948.
- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a theory of mind. *Cognition*, *21*, 37–46.
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175.
- Baron-Cohen, S., Wheelwright, S., Stone, V. E., Jones, R., & Plaisted, K. (1999a). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*(5), 407–418.
- Baron-Cohen, S., Wheelwright, S., Stone, V. E., & Rutherford, M. (1999b). A mathematician, a physicist, and a computer scientist with Asperger syndrome: Performance on folk psychology and folk physics tests. *Neurocase*, *5*(6), 475–483.
- Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on “theory of mind” and cognition. *Brain*, *127*(4), 914–928.
- Blakemore, S. J., Boyer, P., Pachot-Clouchard, M., Meltzoff, A., Segebarth, C., & Decety, J. (2003). The detection of contingency and animacy from

- simple animations in the human brain. *Cerebral Cortex*, 13(8), 837–844.
- Boucher, J., & Lewis, V. (1992). Unfamiliar face recognition in relatively able autistic children. *Journal of Child Psychology & Psychiatry*, 33(5), 843–859.
- Campbell, R., Heywood, C., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia*, 28(11), 1123–1142.
- Carlson, S., & Moses, L. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4), 1032–1053.
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation and play as infancy precursors to language and theory of mind. *Cognitive Development*, 15, 481–498.
- Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences*, 3(3), 115–120.
- Corballis, M. (2003). Recursion as the key to the human mind. In K. Sterelny & J. Fitness (Eds.), *From mating to mentality: Evaluating evolutionary psychology* (pp. 155–171). New York: Psychology Press.
- Cosmides, L., & Tooby, J. (1995). Foreword. In S. Baron-Cohen (Ed.), *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Courchesne, E., Townsend, J., Akshoomoff, N. A., Saitoh, O., Yeung-Courchesne, R., Lincoln, A. J., et al. (1994). Impairment in shifting attention in autistic and cerebellar patients. *Behavioral Neuroscience*, 108(5), 848–865.
- Crichton, M., & Lange-Küttner, C. (1999). Animacy and propulsion in infancy: Tracking, waving, and reaching to self-propelled and induced moving objects. *Developmental Science*, 2(3), 318–324.
- Csibra, G., Biro, S., Koos, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, 27(1), 111–133.
- Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, 28, 479–485.
- Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., et al. (2004a). Early social attention impairments in autism: Social orienting, joint attention and attention to distress. *Developmental Psychology*, 40(2), 271–283.
- Dawson, G., Webb, S. J., Carver, L., Panagiotides, H., & McPartland, J. (2004b). Young children with autism show atypical brain responses to fearful versus neutral facial expressions of emotion. *Developmental Science*, 7(3), 340–359.
- De Villiers, J., & Pyers, J. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief understanding. *Cognitive Development*, 17, 1037–1060.
- Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press.
- Frith, U., Morton, J., & Leslie, A. M. (1991). The cognitive basis of a biological disorder: Autism. *Trends in Neurosciences*, 14(10), 433–438.
- Gerrans, P. (2003). The theory of mind module in evolutionary psychology. *Biology & Philosophy*, 17, 305–321.
- Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, 7(1–2), 145–171.
- Gregory, C., Lough, S., Stone, V. E., Erzincioğlu, S., Martin, L., Baron-Cohen, S., & Hodges, J. (2002). Theory of mind in frontotemporal dementia and Alzheimer's disease: Theoretical and practical implications. *Brain*, 125, 752–764.
- Griffith, E. M., Pennington, B. F., Wehner, E. A., & Rogers, S. (1999). Executive functions in young children with autism. *Child Development*, 70, 817–832.
- Happé, F. (1999). Autism: Cognitive deficit or cognitive style? *Trends in Cognitive Sciences*, 3(6), 216–222.
- Happé, F., Mahli, G. S., & Checkley, S. (2001). Acquired mind-blindness following frontal lobe surgery. A single case study of impaired "theory of mind" in a patient treated with stereotactic anterior capsulotomy. *Neuropsychologia*, 39, 83–90.
- Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59, 771–785.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, 3(1), 80–84.
- Hornak, J., Rolls, E. T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia*, 34(4), 247–261.
- Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain & Cognition*, 44(2), 280–302.
- Keenan, T. (1998). Memory span as a predictor of false belief understanding. *New Zealand Journal of Psychology*, 27(2), 36–43.
- Kumashiro, M., Ishibashi, H., Itakura, S., & Iriki, A. (2002). Bidirectional communication between a Japanese monkey and a human through eye gaze and pointing. *Current Psychology of Cognition*, 21(1), 3–32.
- Leslie, A. M. (1987). Pretence and representation: The origins of "theory of mind". *Psychological Review*, 94(4), 412–426.
- Leslie, A. M. (1992). Pretense, autism, and the theory-of-mind module. *Current Directions in Psychological Science*, 1(1), 18–21.
- Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50(1–3), 211–238.
- Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, 43, 225–251.
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258.

- Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology*, *9*(4), 491–500.
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology & Psychiatry*, *32*, 1081–1105.
- Perner, J. (1991). *The representational theory of mind*. Cambridge, MA: MIT Press.
- Perrett, D., Harries, M., Mistlin, A., Hietanen, J., Benson, P., Bevan, R., et al. (1990). Social signals analyzed at the single cell level: Someone is looking at me, something touched me, something moved! *International Journal of Comparative Psychology*, *4*(1), 25–55.
- Russell, J., Saltmarsh, R., & Hill, E. (1999). What do executive factors contribute to the failure on false belief tasks by children with autism? *Journal of Child Psychology & Psychiatry*, *40*(6) 859–868.
- Rutherford, M., & Rogers, S. J. (2003). Cognitive underpinnings of pretend play in autism. *Journal of Autism & Developmental Disorders*, *33*(3), 289–302.
- Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). The left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, *7*(5), 499–500.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporoparietal junction in “theory of mind. *NeuroImage*, *19*, 1835–1842.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, *55*, 87–124.
- Schultz, R. T. (2005). Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience*, *23*(2–3), 125–141.
- Smith, M., Apperly, I., & White, V. (2003). False belief reasoning and the acquisition of relative clause sentences. *Child Development*, *74*(6), 1709–1719.
- Snowden, J. S., Gibbons, Z., Blackshaw, A., Doubleday, E., Thompson, J., Craufurd, D., et al. (2003). Social cognition in frontotemporal dementia and Huntington's disease. *Neuropsychologia*, *41*, 688–701.
- Sobel, D. M., & Lillard, A. S. (2002). Children's understanding of the mind's involvement in pretense: Do words bend the truth? *Developmental Science*, *5*(1), 87–97.
- Stone, V. E. (2005). Theory of mind and the evolution of social intelligence. In J. Cacioppo (Ed.), *Social neuroscience: People thinking about people*. Cambridge, MA: MIT Press.
- Stone, V. E. (2006). The moral dimensions of human social intelligence. *Philosophical Explorations*, *9*(1), 55–68.
- Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, *10*, 640–656.
- Stone, V. E., & Gerrans, P. (2006). Does the normal brain have a theory of mind? Thoughts on modularity in response to Apperly et al. *Trends in Cognitive Sciences*, *10*(1), 3–4.
- Suddendorf, T. (1999). The rise of the metacognition. In M. C. Corballis & S. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 218–260). London: Oxford University Press.
- Suddendorf, T., & Whiten, A. (2001). Mental evolution and development: Evidence for secondary representation in children, great apes and other animals. *Psychological Bulletin*, *127*(5), 629–650.
- Suddendorf, T., & Whiten, A. (2003). Reinterpreting the mentality of apes. In J. Fitness & K. Sterelny (Eds.), *From mating to mentality: Evaluating evolutionary psychology* (pp. 173–196). New York: Psychology Press.
- Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.
- Wellman, H., & Lagattuta, K. H. (2000). Developing understandings of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd ed., pp. 21–49). Oxford, UK: Oxford University Press.
- Wellman, H., Phillips, A., Dunphy-Lelii, S., & LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science*, *7*(3), 283–288.
- Wellman, H. M. (1988). First steps in the child's theorizing about the mind. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 64–92). New York: Cambridge University Press.
- Woodward, A. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior & Development*, *22*(2), 145–160.