

Early Social Cognition: Alternatives to Implicit Mindreading

Leon de Bruin · Derek Strijbos · Marc Slors

Published online: 1 September 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract According to the BD-model of mindreading, we primarily understand others in terms of beliefs and desires. In this article we review a number of objections against explicit versions of the BD-model, and discuss the prospects of using its implicit counterpart as an explanatory model of early emerging socio-cognitive abilities. Focusing on recent findings on so-called ‘implicit’ false belief understanding, we put forward a number of considerations against the adoption of an implicit BD-model. Finally, we explore a different way to make sense of implicit false belief understanding in terms of keeping track of affordances.

1 Introduction

For several decades the communis opinio (or something near enough) amongst philosophers and psychologists has been that mindreading is a prerequisite for successfully navigating the social world. Mindreading, as Baron-Cohen (2001) puts it, is the ability to ‘infer the full range of mental states (beliefs, desires, intentions, imagination, emotions, etc.) that cause action’, and use these inferences to predict and explain the behavior of others. This allows us to understand that, for example, the soccer player *wants* to play the ball but *unintentionally* tackles the referee, who gives him a red card because he *falsely believes* the player did it on purpose, and so on.

There have been two dominant approaches when it comes to explaining mindreading: theory theory and simulation theory. According to theory theory (TT), the ground rules for mindreading are laid down by a folk psychological theory that specifies how mental states interrelate and inform actions. Simulation theory

L. de Bruin (✉)
Institute for Philosophy II, Ruhr-University of Bochum,
Universitätsstr. 150, 44801 Bochum, Germany
e-mail: lcdebruin@gmail.com

D. Strijbos · M. Slors
Department of Philosophy, Radboud University Nijmegen, Erasmusplein 1, 6500 HD Nijmegen,
The Netherlands

(ST), by contrast, claims that mindreading involves ‘putting ourselves in the other’s shoes’ by simulating his or her mental states. Both TT and ST have been developed in several versions.¹ Nowadays, however, most proponents of TT and ST favor hybrid models that accommodate both theorizing and simulation, although they still argue about the specific parts played by either theory or simulation.²

In this article we will not be concerned with particular articulations of TT, ST or hybrid TT/ST accounts. Nor will we deal with questions about the possible role of theory and/or simulation in a plausible explanation of mindreading. Instead, we take issue with a certain conception of mindreading that is usually presupposed by TT, ST and hybrid accounts. According to this view, mindreading is the primitive and default way of understanding others. As Nichols and Stich (2003) put it, ‘we engage in mindreading for mundane chores, like trying to figure out what the baby wants, what your peers believe about your work, and what your spouse will do if you arrive home late’ (pp. 1–2). Moreover, this conception also has it that mindreading is primarily about employing certain procedures that are rooted in *belief-desire psychology*. In particular, the idea is that mindreading is a powerful social tool because it allows us to manipulate specific combinations of beliefs and desires in order to predict and explain the behavior of others.³ We call this the belief-desire (BD) model of mindreading. In its strongest formulation, the BD-model takes our understanding of others to require the explicit (re)construction of their behavior in terms of a constellation of beliefs and desires, minimally a desire toward some goal and a belief regarding the means.⁴ When predicting the behavior of others, we start with a pair of interlocking beliefs and desires and work our way towards a predicted or anticipated behavioral outcome. In case of behavior explanation, by contrast, we work back from the behavior under consideration to a particular belief-desire pair.

In the next section, we discuss a number of arguments against the conception of mindreading as described above, i.e. against the conjunction of (i) the assumption that mindreading is the default way of understanding others (in terms of both frequency and universality), and (ii) the idea that, in its prototypical form, mindreading results in the attribution of belief-desire pairs. This sets the stage for section three, in which we investigate a popular strategy to counter these criticisms, namely by introducing another distinction: between mindreading as an *explicit* (conscious) manipulation of mental states and mindreading as an *implicit*, sub-personal process (unfolding automatically and without conscious control). This allows proponents of mindreading to preserve both (i) and (ii) at a different level of explanation. However, this move is questionable and perhaps even misguided. In our

¹ There are various versions of TT, for example ‘modular TT’ (Fodor 1992; Leslie et al. 2005), ‘scientific TT’ (Gopnik and Meltzoff 1997), ‘model TT’ (Maibom 2003, Godfrey-Smith 2005), and ‘external TT’ (Braddon-Mitchell and Jackson 2007). ST has also been developed into several directions, such as ‘explicit’ ST (Heal 1986; Goldman 1989), ‘radical’ ST (Gordon 1986, 2008), and ‘implicit’ ST (Hurley 2005, Gallese 2005).

² See, for example, the hybrid ST/TT accounts proposed by Currie and Ravenscroft (2002), Nichols and Stich (2003), Mitchell (2005), Saxe (2005) or Goldman (2006).

³ This conception is taken for granted by many proponents in the mindreading debate. See for example, Currie and Sterelny (2000, p.143), Frith and Happé (1999, p.2) and Scholl and Leslie (1999, p.131).

⁴ These processes obey the principles of belief-desire psychology, chief amongst them being the ‘central action principle’: ‘if A wants P and believes that doing q will bring about p, then ceteris paribus, A will q’ (Borg 2007, p. 6). See Botterill (1996), p. 115, for a similar formulation.

view, the central question that proponents of mindreading need to answer is why implicit forms of social interaction need to be modeled on explicit belief-desire psychology when the latter turns out to be problematic. Are there good reasons to think that the BD-model provides us with the best explanation of our implicit social skills? We think not. Recently, the BD-model has been defended on the basis of findings on so-called ‘implicit’ false belief understanding in infants of 7–25 months. In section four we present a number of considerations against a belief-desire interpretation of these findings, drawing on a critical article by Apperly and Butterfill (2009). Finally, section 5 explores a different way to make sense of implicit false belief understanding in terms of keeping track of affordances.

2 The BD-model of Mindreading: Some Problems

There are basically two ways in which claims about the ubiquity of mindreading can be further specified, namely with respect to its *frequency* (is it used in all our social engagements or only occasionally), and its *universality* (is it a capacity that is shared by all human beings). Both strands have met with severe criticism.

One argument against the frequency-claim can be derived from a simple reflection on the everyday phenomenology of social interaction. If mindreading BD-style is indeed the default procedure by which we come to understand others, then one might expect it to show up in our experience. Goldman (2006), for example, defends a version of ST according to which simulation comes down to a process in which we (a) generate pretend beliefs and desires, (b) feed them into our own offline practical decision-making system, and (c) attribute the resulting pretend decisions to the person we want to understand. Gallagher (2007) has objected that if the simulation procedures prescribed by Goldman are employed in a frequent and explicit fashion, then we should be aware of the different steps that we go through as we consciously simulate the other’s mental states. However, when I interact with others and try to understand them, ‘there is no experiential evidence that I use such conscious (imaginative, introspective) simulation routines’ (2007, p. 65). This is what he calls the ‘simple phenomenological argument’. Similar arguments have been made against explicit versions of TT (e.g., Gallagher 2004; Ratcliffe 2007).

The simple phenomenological argument is perhaps a bit too simple. Goldman (2006) argues that the appeal to phenomenology is problematic because phenomenology is ‘incapable of supporting weighty theses’, hard to agree upon and ‘hotly disputed’ (p. 249). Some have taken this argument a step further. Spaulding (2010) claims that ‘the fallibility of phenomenology is one reason to doubt Gallagher’s phenomenological argument. The total irrelevance of phenomenology is another’ (p. 131). Now this seems to be a bit of an overstatement. It is one thing to say that phenomenology can be mistaken in many instances, but it is quite another thing to claim that it may as well be systematically mistaken—which is what total irrelevance implies. What is more important, however, is the following: it is not clear how an appeal to phenomenology can be conclusive in answering questions about the frequency of BD mindreading. As Gallagher

(forthcoming) himself acknowledges, this rather seems to be an empirical question.⁵ Even if the simple phenomenological argument is not conclusive in this respect, however, one could still maintain that it can be used effectively against those who simply take the dominance of BD-mindreading for granted.

Another argument against the assumed frequency of BD-mindreading is directed at its scientific standing. Advocates of the BD-model often appeal to developmental psychology to support their claims about the importance of mindreading. In particular the false belief task (FBT) has been a popular choice. In the ‘unexpected location’ FBT, for example, children observe a protagonist who sees an object being placed in a certain location (e.g., Wimmer and Perner 1983; Baron-Cohen et al. 1985). The protagonist leaves, and the object is moved. When the protagonist returns, she mistakenly believes the object is still in its initial location. At this point, the child is asked to predict where the protagonist will look for the object. Results show that 3-year-olds fail this task, whereas 4-year-olds typically answer correctly. This has been interpreted as reliably indicating that 3-year-olds do not understand that the protagonist has a false belief about the location of the object, whereas 4-year-olds are capable of distinguishing between how things really are in the world and what other people may falsely believe about such things (e.g., Perner 1991; Gopnik and Wellman 1992; Wellman 2002; Hale and Tager-Flusberg 2003).

It might be argued that an appeal to these findings is not appropriate, because the FBT is not representative for the full scope of social understanding. First, a narrow focus on FBT performance may lead to an overstatement of the role of mindreading in social interaction. Bloom and German (2000), for example, have warned that we are dealing here with an ‘ingenious, but very difficult task that taps (only) one aspect of people’s understanding of the minds of others’ (p. 30). And Gallagher (2004) points out that the FBT is designed to capture a set of very specialized cognitive abilities, which ‘put us in an observational mode and do not capture the fuller picture of how we understand other people’ (p. 204). If this is correct (and we think it is), then this obviously has consequences for the extent to which FBT findings can be used to vindicate claims about the frequency of BD-mindreading. Second, it has been argued that performance on FBT might not amount to a full understanding of the concept of belief. As Hutto (2008) points out, the FBT provides us only with evidence for an isolated understanding of the concept of belief: ‘knowing that children manage to pass false-belief tests, reliably enough, at a certain age under very particular experimental conditions, gives no insight into the extent of their understanding of that concept in other contexts’ (p. 26). This is also emphasized by Carpendale and Lewis (2004, p. 91), who observe that ‘research that explores whether 5-year-olds can use simple false belief knowledge to make inferences about their own and other’s perspectives finds that they singularly fail to do so.’ Third, performance on the FBT does not seem to provide evidence for a mindreading competence that meets the requirements of the BD-model. Folk psychology *stricto sensu*, as Hutto (2008) labels it, at the very least requires the ability to employ belief-

⁵ There are also non-empirical considerations that count against the idea that BD mindreading is ubiquitous, such as the consideration that in many simple daily interactions BD mindreading is not computationally parsimonious (cf. Bermúdez 2003, pp. 31–2). See Spaulding (2010) for a critical discussion of this argument.

desire propositional attitude psychology. In order to make sense of another person's action 'it is not enough to imagine it as being sponsored by a singular kind of propositional attitude; one must also be able to ascribe other kinds of attitudes that act as relevant and necessary partners in motivational crime' (p. 26).⁶

Besides criticizing claims about the ubiquity of belief-desire mindreading in terms of frequency, it is also possible to raise questions about its universality. Several studies have shown that there are cultures whose folk psychologies do not allow for an unambiguous parallel to our belief-desire distinction (e.g., Howell 1981; Lebra 1993; Rosaldo 1980; Vinden 1996; Wellman 1998; Wierzbicka 1992, 2009). Other studies indicate that there might in fact be another, more widespread and possibly dominant form of understanding others, namely, in terms of *socio-situational* factors (e.g., Naito and Koyama 2006; Morris and Peng 1994; Dweck et al. 1995; Al-Zahrani and Kaplowitz 1993; Miller 1984; Straus 1977; Briggs 1970; see Lillard 1998 for a systematic review). These findings suggest that the tendency to employ the BD-model is specific to European American cultures, and may be in part a cultural acquisition (Beauvieux and Dubois 1988).

What cross-cultural studies also show is that there is large variability in performance on the FBT.⁷ As we mentioned earlier, it has traditionally been taken for granted that children typically pass this task around 4 years of age. However, over the last decade various experiments have casted doubt on this assumption. For example, in a meta-analysis comparing non-Western (China and Hong Kong) and North American (United States and Canada) children, Liu et al. (2008) found large variations in the onset of FBT performance.⁸ These variations are even larger when non-Western, non-industrialized countries are included (e.g., Vinden 1996, 1999, 2002).⁹

Recently, Kobayashi and Temple (2009) have shown that the cross-cultural differences found in FBT performance have their counterpart at the neurobiological level. The investigators conducted a meta-analysis in which they compared neuroimaging studies that employed the FBT in American, British, French, German, Japanese and Swedish culture. The results showed considerable cross-cultural differences in the activation of brain regions such as the temporal pole (TP) and the temporo-parietal junction (TPJ), which might reflect cross-cultural variability at the neurobiological level. According to Kobayashi and Temple (2009), the diminished

⁶ Several developmental theorists claim that children acquire a concept of a desire-like state a year or so before they perform successfully on the FBT (e.g., Wellman 1990; Wellman and Phillips 2001), and there is empirical support for such an asymmetry in acquisition as well (e.g., Repacholi and Gopnik 1997; Rakoczy et al. 2006). But Hutto's point above still holds, one might argue, to the extent that understanding (false) belief and desire in isolation does not suffice for mindreading BD-style.

⁷ Notice that the limits of the FBT (as discussed in the previous section) might still leave open the possibility that people from different cultures have an understanding of belief that is not measured by the FBT. At the same time, however, this is something that still has to be (empirically) demonstrated by proponents of the BD-model. We thank an anonymous reviewer for this point.

⁸ For example, Canadian children acquire false belief understanding around 3 years of age, whereas Hong Kong children showed similar performance as much as 2 years later. And Naito (2003) reports that 40% of 5–6 years old Japanese children still fail the standard FBT (see also Koyasu 1997; Goushiki 1999).

⁹ Vinden (1996), for example, investigated false belief understanding among 4 to 8-years-old Junin Quechua children in Peru, and discovered that the majority of the oldest children were unable to respond correctly to standard false belief questions. Moreover, she did not find evidence for an improvement of children's performance with increasing age.

activity in TPJ in Japanese children and adults indicates a reduced dependency on explanatory strategies that refer to individual beliefs and desires. This is in line with studies by Naito (2007) and Naito and Koyama (2006), who found a preference in Japanese culture for explaining behavior in terms of social-situational factors instead of individual mental states.

These cross-cultural differences in mature folk psychological vocabularies, explanation/prediction strategies, and performance on the FBT are not incompatible per se with the BD-model of mindreading. Nichols and Stich (2003), for example, recognize that the cross-cultural differences mentioned above pose problems for the BD-model. However, they think these problems can be solved by conceiving of beliefs and desires as ‘thin’ psychological states that are universally shared, unlike the inferentially rich and sophisticated ‘thick’ psychological states that are only exploited in the mature folk psychologies of specific cultures. Nichols & Stich propose to talk about the technical terms ‘bel’ and ‘des’ in place of ‘belief’ and ‘desire’, because ‘the concepts produced by the detection mechanisms really are quite thin - they do not come loaded with a rich set of inferential connections’ (p. 206). The question is what such ‘thin’ concepts precisely amount to—which core features of the ‘thick’ propositional attitudes remain applicable. If these turn out to be surprisingly few, then we need to ask ourselves what we gain by framing cross-culturally shared social skills in terms of a BD-model. As we show below, this worry also arises when we try to employ the BD-model in order to account for early socio-cognitive development in infants.

3 ‘Implicit’ Mindreading?

The arguments discussed so far are certainly not decisive for our case against the BD-model. Nevertheless, it is interesting to see how proponents of the BD-model respond to them. As far as the argument from phenomenology is concerned, advocates of TT, ST or hybrid TT/ST usually point out that we do not use the BD-model in order to theorize about or simulate others in a conscious and explicit way. If theorizing or simulating is an unconscious and implicit process, then what we experience or seemingly experience is not a good guide for what is ‘really’ happening in such cases, and the appeal to phenomenology is inappropriate. As Herschbach (2008) argues, ‘phenomenological claims have bite at the personal level (...). [A]ppeals to phenomenology and other arguments do not, however, rule out theory theory and simulation theory as accounts of the subpersonal processes underlying social [cognition]’ (p. 223). Spaulding (2010) makes a similar point when she writes that ‘with mindreading, there is a process (theorizing or simulating), and there is a product (an explanation or a prediction). In general, neither the process nor the product need be consciously accessible, let alone phenomenologically transparent’ (p. 131).

However, on what grounds and by which criteria do we establish that this kind of ‘implicit’ mindreading is indeed best understood in terms of the BD-model? Proponents of TT, ST and hybrid TT/ST often argue that the BD-model can be vindicated if we pay attention to important developmental precursors to explicit mindreading, like the emergence of pretend play abilities (Leslie 1987; Perner 1991;

Lillard 2002) and capacities such as intention detection, eye-direction detection, and shared attention (Baron-Cohen 1995). Interestingly enough, a similar argument is used to defend the universality of the BD-model against the cross-cultural evidence discussed in the previous section. Many proponents of TT in particular (e.g., Wellman 1998; Gauvain 1998; Scholl and Leslie 1999) argue that the cross-cultural differences mentioned by Lillard (1998) and others are not problematic for the BD-model as long as we distinguish between ‘opulent’ and ‘core’ accounts of mindreading. According to opulent accounts, mindreading employs a relatively wide-ranging and complex set of folk psychological concepts (cf. Churchland and Churchland 1998). Core accounts, on the other hand, argue that mindreading is restricted to a much smaller subset of core concepts (cf. Scholl and Leslie 1999). Although core accounts typically do not deny that mindreading properly denotes the rich and complex conceptual framework addressed by opulent accounts, they focus on the alleged *core* of this conceptual framework. This allows them to dismiss the cross-cultural findings on mindreading as irrelevant for their position. Wellman (1998), for example, concludes that, despite the fact that our mindreading capacities might be ‘quite different from one another worldwide’, they could still be grounded in ‘the initial framework assumptions of young children’ (p. 35). And Scholl and Leslie (1999) also maintain that the cross-cultural differences pertain only to ‘the full-fledged mature ToM [Theory of Mind] competence, rather than to its general character in early acquisition’ (10).¹⁰

In other words, the BD-model is thought to be implicit insofar as it is applicable as an explanatory model for early developing socio-cognitive capacities. But there are also other ways to capture the implicitness of the BD-model. It is sometimes claimed that the BD-model is implicit in the sense that the logical principles of belief-desire psychology can be directly found at the *neurobiological* level. According to Goldman’s (2006) hybrid TT/ST account, for example, explicit mindreading is executed by simulation routines, which are in turn implemented by an underlying implicit theory. ‘How could simulation be executed unless an algorithm for its execution is tacitly represented at some level in the brain? Isn’t such an algorithm a sort of theory?’ (p. 33). Proposals of this sort usually go hand in hand with claims about the innateness of BD-psychology. Scholl and Leslie (1999), for instance, back up their core BD-account of mindreading by arguing for the existence of a universal innate module (a ‘Theory-of-Mind-Mechanism’), which contains the basic meta-representational concepts belief, pretense and desire.

The fact that claims about the innateness of the BD-model and its virtues as an explanatory model of early developing mindreading skills often come in tandem is not surprising if we consider that the former are usually supposed to do some explanatory work with respect to the latter. The attractiveness of the BD-model as an

¹⁰ Wellman (1998) agrees with Scholl and Leslie (1999) about the existence of a core framework for mindreading. However, he thinks that such a core account will be different from the BD-model, since ‘belief, desire, and emotion are English terms with English specifics, unlikely to properly capture core conceptual notions. [...] Terms akin to think, want and feel [...] may be common to all the world’s languages but certainly not belief, desire and emotion’ (p. 34). If Wellman’s observation is correct, then TT and ST proponents of the BD-model are on the wrong track, since they assume that a core account of mindreading should at least include beliefs and desires. See Mills (2001) for further discussion of the distinction between opulent and core accounts of mindreading.

explanatory model of early socio-cognitive capacities is intimately intertwined with the idea that ‘humans everywhere interpret the behavior of others in [...] mentalistic terms because we all come equipped with a “theory of mind” module [...] that is compelled to interpret others this way, with mentalistic terms as its natural language’ (Tooby and Cosmides 1995, p. xvii). If human beings are indeed born with an innate module that compels us to understand others in terms of the BD-model, then it also seems that we do not have to be explicitly aware of this when we interact with others.

For each of these interpretations of implicit mindreading, it is important to investigate the merits of the BD-model as a suitable explanatory model. In what follows, however, we will concentrate on the pivotal question why the BD-model should be used as an explanatory model for social-cognitive skills that emerge early in development. More in particular, we present advocates of implicit BD-mindreading with the following challenge: (i) if the BD-model was originally postulated as an explanatory model of how we exercise our mindreading skills in an explicit fashion (and we take it this is a reasonable assumption), and (ii) it now turns out that the explicit BD-model is problematic (given the objections discussed in section 2), then (iii) why should we model our account of implicit mindreading on the BD-model as it was explicitly construed? Why should we expect significant explanatory pay-off by modeling our early social abilities on an explicit BD-model that is already problematic?

Notice that our worry is not that the BD-model *could not* be used as a model of early developing socio-cognitive abilities. The issue is rather why we *should* do so—why it would be our *best* choice. This question becomes particularly pressing when we realize that the BD-model does not explain the difference between implicit and explicit forms of mindreading. Instead, it requires us to introduce new explanatory factors to do the job—factors that are not derived from the BD-model itself. In this light, reconsider Nichols and Stich’s (2003) ‘thin’ rendering of the BD-model. How ‘thin’ may the concepts of belief and desire become while still remaining recognizable as such? Which core features of the propositional attitudes do they need to retain in order to maintain their explanatory power? In the next section, we will discuss some disadvantages of the BD-model as an explanatory model of early developing socio-cognitive abilities. We will focus in particular on recent findings on so-called ‘implicit’ false belief understanding.

4 The BD-model as an Explanatory Model of Early Socio-cognitive Abilities

The decisive importance of early socio-cognitive abilities in the mindreading debate has also been emphasized by proponents of enactivism (e.g., Hutto 2004, 2008; Reddy and Morris 2004; Iacoboni and Dapretto 2006; Ratcliffe 2007; Gallagher and Zahavi 2008; Fuchs and De Jaegher 2009). Their main message has been the exact opposite of the one preached by proponents of the implicit BD-model. According to enactivism, most of our encounters with others can be explained in terms of embodied practices that do not require belief-desire psychology. These embodied practices are said to be *primary* to BD-style mindreading, in the sense that they involve capabilities that come earlier in development and are likely to be partially

innate.¹¹ Furthermore, they are also said to be primary in the sense that they *continue* to characterize most of our social engagements and remain the *default* mode of how we understand others.

These claims have been hotly contested recently. For example, Herschbach (2008), Spaulding (2010) and Jacob (2011) argue that it has yet to be shown that BD-model mindreading accounts are explanatory inferior to recent ‘embodied cognition’ (Spaulding’s terminology) or enactivist theories. Although this might be true for the plausibility of the BD-model in general, there are several reasons why we should not expect too much from applying an implicit version of the BD-model to early developing socio-cognitive abilities.

To start with, it seems to be somewhat ad hoc to adopt an implicit BD-model solely in response to the failure of its explicit counterpart. If the BD-model is not representative for the surface properties of mindreading, then we need an independent rationale for adopting it as an explanatory model to characterize early social development. As Gallagher (2005) points out, it is not sufficient to appeal to the classic FBT (as discussed in section 2), since in this test children are asked to give an explicit prediction of another agent’s behavior on the basis of his or her false belief. The classic FBT is an ‘elicited response’ false belief task or ‘ER-FBT’ (Baillargeon et al. 2010), which involves conscious processing and does not address BD-psychology operating on an implicit level.

Recently, however, the BD-model has been defended on the basis of ‘spontaneous response’ false belief tasks or ‘SR-FTBs’, which no longer require an explicit answer to a question about the protagonist’s belief. Instead, children’s understanding of false belief is inferred from the behavior they spontaneously produce (cf. Baillargeon et al. 2010). Employing ‘violation of expectation’ and ‘anticipatory looking’ paradigms, several studies have claimed that false belief understanding emerges at a considerably earlier age, in 25-month-olds (Southgate et al. 2007), 15-month-olds (Onishi and Baillargeon 2005), 13-month-olds (Surian et al. 2007), and even 7-month-olds (Kovács et al. 2010).

Onishi and Baillargeon (2005), for example, conducted a violation of expectation experiment in which 15-month-old infants were familiarized with a protagonist hiding a toy in one of two locations. The protagonist left, and the toy was moved without her knowledge. Then the infants were shown scenes of the protagonist searching for the hidden toy, either where she falsely believes it to be, or where it was actually located. The experimenters found that infants looked significantly longer at those scenes in which the protagonist searched at the correct location despite her false belief about where the toy was hidden. They concluded that this measure of action anticipation in fact demonstrated an early understanding of false belief.

Some have labeled the early manifestation of false belief understanding ‘implicit’ (Ruffman et al. 2001), because the infants do not seem to be explicitly aware of the false belief of the protagonist. That is, although the infants’ behavioral responses indicate that they are sensitive to the protagonist’s false belief, they give an incorrect answer when asked a direct question about it (cf. Clement & Perner 1994).

¹¹ Gallagher (2001, p. 91), for example, claims that ‘before we are in a position to [...] simulate, explain or predict mental states in others, we are already in a position to interact with and to understand others in terms of their gestures, intentions and emotions, and in terms of what they see, what they do or pretend’.

However, the crucial question is whether we have to interpret the above results in terms of the BD-model. What do we gain by interpreting these findings as instances of false belief understanding? To talk sensibly about belief ascription in this context, at least the following criteria need to be met: (i) beliefs have propositional content, (ii) beliefs are representational states, i.e. they capture the way the agent (mis) represents the world, and (iii) beliefs have a specific direction of fit and play a psychological role in the cognitive economy accordingly.

With respect to (i): it is not at all clear why explanatory models of early developing socio-cognitive abilities would have to be limited by this constraint. In fact, most current models that try to characterize implicit false belief understanding are *not* limited in this sense. Baillargeon et al. (2010), for example, recently proposed a modular theory theory account according to which infants come equipped with an innate psychological-reasoning system that consists of two sub-systems: sub-system 1 and sub-system 2. Sub-system 1 enables infants to attribute motivational states and ‘reality-congruent informational states’ to other agents, and is well in place by the end of the first year of life.¹² Sub-system 2 is more advanced in the sense that it also enables infants to attribute ‘reality-incongruent informational states’ to other agents, and becomes operational in the second year (cf. Scott and Baillargeon 2009). Baillargeon et al. (2010) argue that the spontaneous response or SR-FBT findings are indeed indicative of false belief understanding, and they explain this in terms of sub-system 2 processes. At the same time, however, they merely characterize this understanding as the ability to attribute ‘reality incongruent informational states’.

This obviously falls short of the much more advanced capacity to attribute *propositional* states to others. As Apperly and Butterfill (2009) have recently pointed out, ‘in terms of content [...] no study has yet suggested that infants track beliefs involving both the features and location of an object (e.g., ‘The red ball is in the cupboard’) or that they track beliefs whose contexts can be represented only using quantifiers (e.g., ‘there is no red ball in the cupboard’); or that, in tracking beliefs, they are sensitive to modes of presentation’ (p. 957). Apperly & Butterfill argue that SR-FBTs only require that infants can track attitudes to objects’ locations.

With respect to (ii): the SR-FBT shows that infants anticipate agents to reach for a non-actual location. But such anticipation falls short of evaluating them as agents who have a view on the world that might be in conflict with how things actually are. There are two sides to this requirement. The first one is closely tied to the first criterion above: in order for a belief-like state to represent a certain state of affairs, it must have truth conditions. Only states with propositional content can be judged true or false. Secondly, in order to understand an agent’s mental states as (mis) representations of the environment, one must be sensitive to the mode of presentation under which the relevant environmental features are made present to the agent. That is, one must be able to distinguish between what someone represents and how she represents it. This capacity for level 2 perspective taking is not required

¹² Reality-congruent informational states specify what knowledge or accurate information the agent possesses about the scene. Motivational states, by contrast, are defined as states that specify the agent’s motivation in the scene and include goals and dispositions.

for successful performance on the SR-FBT. Keeping track of an agent's attending to an object falls short of understanding her as believing *of* an object at location A *that* it is at location B.

In this context it is interesting to consider Apperly & Butterfill's alternative interpretation of infants' performance on the SR-FBT. On their interpretation, infants are sensitive to the agent's belief only insofar as it *registers* the object. First, however, Apperly & Butterfill explain the simpler notion of *encountering*. Encountering is defined as 'a relation between an individual, an object and a location, such that the relation obtains when the object is in the individual's field' (p. 962). A field is defined, simply, as a certain region of space around the individual. Building on this, registering is defined as a slightly more complex psychological relation that obtains between an individual, an object and a location. An individual is said to register an object at a location when (a) she encounters the object at the location and (b) has not since encountered it somewhere else. A registering is off target when the object registered is not located where it is registered to be. The importance of the concept lies in its connection to action: 'One can understand registration as an enabling condition for action, so that registering an object and location enables one to act on it later [...] Further, registration also can be understood as determining which location an individual will direct their actions to when attempting to act on that object' (962). Tracking an agent's registration of something does not require sensitivity to her mental states as representings *as*. In fact Apperly & Butterfill predict failure on level 2 perspective taking to be among the signature limits imposed by sensitivity to mere registrations.

With respect to (iii): we should consider the possibility that implicit social understanding does not respect the clear-cut divide between states with a mind-to-world direction of fit (prototypically, beliefs) and states with a world-to-mind direction of fit (prototypically, desires). Apperly & Butterfill's registerings are presented as *belief*-like states, with a mind-to-world direction of fit. It is, however, possible and even plausible to interpret registerings and the SR-FBT results without the restriction of a specific direction of fit. Registerings are defined in terms of encounterings, and they inherit their direction of fit from the (mind-to-world) direction of fit that the notion of 'encountering' seems to imply. Yet, encounterings only have a mind-to-world direction of fit on a purely *passive, observational* understanding of them. In view of developments in the philosophy and neuroscience of perception such as enactivist and ecological theories of perception (e.g. Gibson 1979; Hurley 1998, Noë 2004), however, it seems much more plausible to conceive of encountering in terms of an agent's being sensitive to the specific affordances of the object-at-a-location encountered. The agent's 'field' in which the object occurs, then, is primarily a field of *affordances* (cf. Rietveld 2008). Such an enactivist conception of encountering does not imply a direction of fit, precisely because Gibsonian affordances have neither a mind-to-world, nor a world-to-mind direction of fit. If registering is defined in terms of an enactivist notion of encountering, then, registering does not have a specific direction of fit either. Hence, on an enactivist view it is simply incorrect to characterize the registerings in terms of which we can understand the SR-FBT results in a way that is modeled on BD mindreading.

5 An Alternative Reading of Spontaneous Response FBT Results

The above considerations show that the SR-FBT findings fail to meet the requirements for an interpretation in line with the BD-model of mindreading. Moreover, insisting on such an interpretation gives rise to a developmental paradox: if infants' performance on the SR-FBT is representative for an understanding of false belief, then why do they fail the ER-FBT? Since this paradox is a direct consequence of the BD-model, it cannot be dealt with by appeal to belief-desire psychology. Baillargeon et al. (2010) provide us with a good illustration of why this doesn't work. Observing that we need a careful analysis of the task requirements of SR- and ER-FBTs, they propose that the SR-FBT only involves (i) a process of false belief representation, whereas the ER-FBT also requires (ii) a response selection process (when asked the test question, children must access their representation of the agent's false belief to select a response) and (iii) a response-inhibition process (when selecting a response, children must inhibit any prepotent tendency to answer the test question based on their own knowledge; cf. Scott and Baillargeon 2009).

What is problematic about this proposal is that it seems arbitrary to argue that spontaneous response versions of the FBT only require false belief representation, whereas elicited response versions also require selection processing and response-inhibition. If we accept that the SR-FBT involves false belief representation, then it is not clear why it does not require selection processing and response-inhibition as well. For infants still have to select a false belief among other beliefs, and in order to do so they have to inhibit their default attribution of true beliefs.¹³ And if this is correct, then it does not work to argue that failure on the ER-FBT is due to the joint activation of false-belief-representation processes and response-selection processes, which 'overwhelms' the child's limited information-processing resources. Of course, it might still be true that ER-FBTs require more information processing resources. But when these explanations are put to work in the service of an implicit BD-model, they strike us as somewhat far-fetched.

Apperly & Butterfill's notion of registering seems to be much more promising when it comes to characterizing implicit false belief understanding. It is not only more fine-grained than, for example, Nichols and Stich's (2003) technical counterpart of belief ('bel'), but modeling the infants' performance on the SR-FBT in terms of registering also places specific limitations on their behavior. Apperly and Butterfill (2009) mention a few such 'signature limits': registering does not permit tracking mental states that involve quantifiers or complex combinations of properties of objects (e.g. their location and color), and it does not support level 2 perspective taking, i.e. keeping track of modes of presentation. Importantly, such cognitive limitations are arbitrary from the point of view of the BD-model. There is nothing in the concept of belief or desire that would predict them; any explanation of these restrictions on socio-cognitive performance has to be imported from outside the BD-model itself.

However, the question is whether the term 'registering' is sufficiently demarcated from the folk psychological concept of belief. Apperly and Butterfill (2009) still

¹³ Baillargeon et al. (2010) assume that infants by default attribute true beliefs to other agents. See Leslie (2000) or Nichols and Stich (2003) for the rationale behind this assumption.

hold that children have the ‘ability to ascribe simple forms of mental content, at least in the form of belief-like states’ (2009, p. 965). But in view of the observations made in the previous section, it is not at all clear that SR-FBT performance requires the ability to process belief-like states with a mind-to-world direction of fit. ‘Encountering’ may just as well be conceived in terms of sensitivity to the affordances of objects.

In what follows we therefore present an alternative, enactivist version of what Apperly & Butterfill call ‘registering’ - one that provides us with a plausible interpretation of what goes on in the SR-FBT but does not involve the ascription of belief-like states, however minimally defined. The difference with Apperly & Butterfill’s proposal is that on our account, infants do not track relations between agents and objects-at-locations *simpliciter*. Instead, they track relations between objects-at-locations and *the actions that these objects afford*. What qualifies this as *social cognition* is precisely the fact that infants track affordances *for other people*.

The introduction of the concept of affordances in this context is crucial, because it helps us to avoid thinking of the infant as postulating an intermediate factor between a perceived object and the predicted actions of an observed agent. Belief-like states such as registrations are internal representations mediating between a perceived object and an agent’s actions on that object. However, when infants are thought of as perceiving an array of possible actions an object affords to some agent, there is no need for processing internal representations of this kind.

Consider, for instance, the following affordance-based interpretation of what happens in violation of expectation experiments (e.g., Onishi and Baillargeon 2005; Surian et al. 2007). First, the infant perceives the affordances of an object-at-a-location (e.g. a ball in a box) for an observed agent. This results in the implicit ascription of an array of possible actions to that agent (where actions are individuated in terms of location, given that the graspable object is invisible). The probability of these actions depends on what has happened during the familiarization trials, and what else will happen in the overall scenario. In this case, the infant will select *reaching for a location* as the most salient affordance for the agent on the basis of the agent’s behavior during the familiarization trials. During the test trial the infant notices that the affordances are changed (e.g. the ball is put into the other box). Yet it is also aware of the fact that the agent looks away or is absent, so that there is no reason to update the *reaching for a location* affordance. In other words, the behavioral disposition the infant ascribes to the agent remains stable across experimental conditions. This explains why the infant’s expectation is violated when the agent undertakes a different course of action. A similar case could be made for what happens in anticipatory looking experiments.

Importantly, although our interpretation does not draw on the ascription of beliefs or belief-like states, it does presuppose a rich notion of what it is to perceive an affordance. Gallagher (2008) has shown how such a rich notion of perception, what he calls ‘smart perception’, obviates the need for attributing beliefs or belief-like states. Of course, this is not to deny that there is internal processing going on during the SR-FBT. The infant’s visual cortex is continuously processing information about the scene, i.e. the interactions between the agent and the object, in a distributed fashion. But this does not involve representations with propositional content and a specific direction of fit.

Take the ‘active helping’ SR-FBT by Buttelmann et al. (2009), for instance, which investigated whether 18-months-old infants are able to actively assist the experimenter in a false belief situation. In this experiment, infants watched as a toy was transferred from box A to box B while an experimenter either witnessed the transfer (true belief condition) or not (false belief condition). Then the experimenter attempted unsuccessfully to open box A—the empty box. In the true belief condition, infants could follow their natural tendency to help the experimenter by opening box A for him. In the false belief condition, if infants understood the experimenter’s false belief, they had to infer that he wanted the toy he thought was in there. In this case they should not simply go help him open box A, but rather go to box B and extract the toy for him. The results indicated that, by 18 months of age, infants were able to do that.

An affordance-based explanation of the active helping behavior in this experiment starts with the infant’s perception of goal-directedness in the agent’s behavior during the familiarization trials. Visual habituation studies indicate that this capacity is already within the reach of 5-month-olds, who respond selectively to the goals of other agents rather than the physical details of their actions (e.g., Biro and Leslie 2007; Gergely and Csibra 2003; Woodward 1998, 2005). On the basis of her perception of the agent’s goal-directed behavior, the infant selects *reaching for location A* as the most salient affordance for the agent. Then, the infant notices that the object is moved to box B. Depending on whether the agent is present or absent during the transfer of the object, she will subsequently update the affordance for the agent into either *reaching for location A* (true belief condition) or *reaching for location B* (false belief condition).

According to such an affordance-based explanation, passing the SR-FBT requires infants to notice that an observed change in affordance does (or does not) imply a change of the agent’s future action if the change is (or is not) visually or otherwise accessible to that agent. In order to do this successfully, infants need to be able to (i) deal with the perceptual incongruencies between their own perspective and that of the other agent, and (ii) distinguish the actions an object affords for themselves from those it affords for another agent. This requires various executive functions, such as response inhibition (e.g., Perner & Lang 1999) and working memory (Perner et al. 2002). Again, however, there is no need for processing representations with propositional content and a specific direction of fit.

One of the advantages of an affordance-based approach to SR-FBTs is that it brings out the fact that these tasks require infants to perceive and process different sorts of information. For example: during the familiarization trials in the study by Onishi and Baillargeon (2005), infants select the relevant affordances for the agent on the basis of her reaching for a location. In the study by Southgate et al. (2007), by contrast, infants do so on the basis of the agent’s *looking behavior* (even though the agent wears a visor, the infants are able to perceive what entities are in the agent’s visual field). Such a further qualification of what is exactly perceived during the various stages of SR-FBTs makes it possible to differentiate between them in terms of processing requirements.

So far, most SR-FBTs have investigated the infant’s ability to anticipate action affordances of other agents on the basis of the *location* of a given object. However, recently some of them have also addressed the infant’s capacity to take into account

the object's *identity* (e.g., Song & Baillargeon 2008, Scott and Baillargeon 2009). Baillargeon et al. (2010) have argued that the results of these experiments, in particular those found by Song & Baillargeon (2008), necessitate an interpretation in terms of false belief. In this study, 14.5 month old infants witnessed an experimenter's hand placing a doll with blue hair and a stuffed skunk with a pink bow on placemats or in shallow containers during familiarization trials. Afterwards, an agent showed her preference for either of the two toys by reaching for it. Then, without the agent watching, the toys were put in two boxes, one of which had a tuft of blue hair on the lid, suggesting it contained the doll. The doll was placed in the plain box, and the skunk in the box with the tuft of blue hair on it. When the agent returned and reached for the box with the tuft of hair after having showed a preference for the skunk, the infants looked considerably longer than when she reached for the plain box—and conversely.

These results might pose difficulties for the association interpretation of the SR-FBT offered by Perner and Ruffman (2005), but they are readily explained in terms of the infant's perception of affordances. Both doll and skunk afford grasping actions, but in the familiarization trials the probability of grasping one toy rather than the other is established. The blue tuft of hair on the box in the test trials is unreflectively perceived as an indicator of the presence of the doll and therefore as affording doll-related actions. Suppose the agent showed preference for the doll in the familiarization trial. Since the agent has not witnessed the change of location of the toys, the infant will perceive the box with the tuft of hair as providing the most salient grasping affordance for the agent during the test trial. The infant will therefore expect the agent to look in that box. This expectation will then be violated when she reaches for the plain box.

Although these affordance-based interpretations are still sketchy, we hope to have demonstrated that they are at least feasible and offer a promising way to think about the results of SR-FBTs. As we have shown in section 4, there are several reasons not to interpret the SR-FBT findings in terms of the BD-model. However, modeling early social cognitive capacities in terms of associations between agents, objects and locations seems to reduce the infant's perception of other people to the perception of moving objects. Much of the resistance to Perner and Ruffman's (2005) proposals in this direction may be grounded precisely in this fact. It is quite generally felt that infants treat other people as more than moving objects; they treat them as intentional systems in their own right, as 'other minds'. One way to meet this intuition is to articulate a notion of association that goes beyond merely perceiving spatial-temporal contiguity between agent and object. In this article, however, we have proposed a different approach: infants treat other people as minded beings in the sense that they treat them as appropriate responders to affordances. The capacity to perceive affordances also explains their early sensitivity to false beliefs.

References

- Al-Zahrani, S., and S. Kaplowitz. 1993. Attributional biases in individualistic and collectivistic cultures. *Journal of Personality and Social Psychology* 47: 793–804.
- Apperly, I., and S. Butterfill. 2009. Do humans have two systems to track beliefs and belief-like states? *Psychological Review* 116: 953–970.
- Baillargeon, R., R.M. Scott, and H. Zijing. 2010. False-belief understanding in infants. *Trends in Cognitive Science* 14(3): 110–118.
- Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind*. MIT Press/Bradford Books.
- Baron-Cohen, S. 2001. Theory of mind in normal development and autism. *Prisme* 34: 174–183.
- Baron-Cohen, S., A.M. Leslie, and U. Frith. 1985. Does the autistic child have a “theory of mind”? *Cognition* 21: 37–46.
- Beauvieux, J.L., and N. Dubois. 1988. The norm of internality in the explanation of psychological events. *European Journal of Social Psychology* 18: 299–316.
- Bermúdez, J.L. 2003. The domain of folk psychology. In *Minds and Persons*, ed. A. O’Hear, 25–48. Cambridge: Cambridge University Press.
- Biro, S., and A.M. Leslie. 2007. Infants’ perception of goal-directed actions. Development through cues-based bootstrapping. *Developmental Science* 10(3): 379–398.
- Bloom, P., and T. German. 2000. Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77: B25–B31.
- Borg, E. 2007. If mirror neurons are the answer, what was the question? *Journal of Consciousness Studies* 14: 5–19.
- Botterill, G. 1996. Folk psychology and theoretical status. In *Theories of Theories of Mind*, ed. P. Carruthers and P. Smith, 184–199. Cambridge: Cambridge University Press.
- Braddon-Mitchell, D., and F. Jackson. 2007. *Philosophy of mind and cognition*. Oxford: Blackwell.
- Briggs, J. 1970. *Never in anger: a portrait of an Eskimo family*. Cambridge: Harvard University Press.
- Buttelmann, D., M. Carpenter, and M. Tomasello. 2009. Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112: 337–342.
- Carpendale, J.L.M., and C. Lewis. 2004. Constructing and understanding the mind: the development of children’s social understanding within social interaction. *The Behavioral and Brain Sciences* 79: 79–151.
- Churchland, P.M., and P.S. Churchland. 1998. *On the contrary: critical essays 1987–1997*. Cambridge: MIT Press.
- Clement, W.A., and J. Perner. 1994. Implicit understanding of belief. *Cognitive Development* 9: 377–395.
- Currie, G., and I. Ravenscroft. 2002. *Recreative minds: imagination in philosophy and psychology*. New York: Oxford University Press.
- Currie, G., and K. Sterelny. 2000. How to think about the modularity of mind-reading. *The Philosophical Quarterly* 50: 145–160.
- Dweck, C.S., C. Chiu, and Y.Y. Hong. 1995. Implicit theories and their role in judgments and reactions: a word from two perspectives. *Psychological Inquiry* 6: 267–285.
- Fodor, J. 1992. A theory of the child’s theory of mind. *Cognition* 44: 283–296.
- Frith, U., and F. Happé. 1999. Theory of mind and self consciousness: what is it like to be autistic? *Mind & Language* 14: 1–22.
- Fuchs, T., and H. De Jaegher. 2009. Enactive intersubjectivity: participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences* 8(4): 465–486.
- Gallagher, S. 2001. The practice of mind: theory, simulation, or interaction? *Journal of Consciousness Studies* 8: 83–107.
- Gallagher, S. 2004. Understanding interpersonal problems in autism: interaction theory as an alternative to theory of mind. *Philosophy, Psychiatry, and Psychology* 11: 199–217.
- Gallagher, S. 2005. *How the body shapes the mind*. New York: Oxford University Press.
- Gallagher, S. 2007. Simulation trouble. *Social Neuroscience* 2: 353–365.
- Gallagher, S. 2008. Direct perception in the intersubjective context. *Consciousness and Cognition* 17: 535–543.
- Gallagher, S. (forthcoming). In defense of an embodied approach to social cognition.
- Gallagher, S., and D. Zahavi. 2008. *The phenomenological mind: an introduction to philosophy of mind and cognitive science*. London: Routledge.
- Gallese, V. 2005. ‘Being like me’: self-other identity, mirror neurons and empathy. In *Perspectives on Imitation: From Cognitive Neuroscience to Social Science: Mechanisms of Imitation and Imitation in Animals*, ed. S. Hurley and N. Chater, 101–118. Cambridge: MIT Press.

- Gauvain, M. 1998. Culture, development, and the theory of mind: comment on Lillard (1998). *Psychological Bulletin* 123: 37–42.
- Gergely, G., and G. Csibra. 2003. Teleological reasoning in infancy: the naive theory of rational action. *Trends in Cognitive Sciences* 7: 287–292.
- Gibson, J.J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Godfrey-Smith, P. 2005. Folk psychology as a model. *Philosopher's Imprint* 5: 1–16.
- Goldman, A. 1989. Interpretation psychologized. *Mind & Language* 4: 161–185.
- Goldman, A. 2006. *Simulating minds: the philosophy, psychology and neuroscience of mindreading*. New York: Oxford University Press.
- Gopnik, A., and A.N. Meltzoff. 1997. *Words, thoughts, and theories*. Cambridge: MIT Press.
- Gopnik, A., and H. Wellman. 1992. Why the child's theory of mind really is a theory. *Mind & Language* 7: 145–172.
- Gordon, R.M. 1986. Folk psychology as simulation. *Mind & Language* 1: 158–171.
- Gordon, R.M. 2008. Beyond mindreading. *Philosophical Explorations* 11(3): 219–222.
- Goushiki, T. 1999. Preschool children's understanding of self-mind and other-mind using 'theory of mind' tasks. *Japanese Journal of Educational Psychology* 47: 354–363.
- Hale, C.M., and H. Tager-Flusberg. 2003. The influence of language on theory of mind: a training study. *Developmental Science* 6(3): 346–359.
- Heal, J. 1986. Replication and functionalism. In *Language, Mind, and Logic*, ed. J. Butterfield, 135–150. Cambridge: Cambridge University Press.
- Herschbach, M. 2008. Folk psychological and phenomenological accounts of social perception. *Philosophical Explorations* 11(3): 223–235.
- Howell, S. 1981. Rules not words. In *Indigenous Psychologies*, ed. P. Heelas and A. Lock, 133–143. London: Academic.
- Hurley, S.L. 1998. *Consciousness in action*. Cambridge: Harvard University Press.
- Hurley, S. 2005. Active perception and perceiving action: The shared circuits hypothesis. In *Perceptual Experience*, eds. T. Gendler and J. Hawthorne, 205–259. New York: Oxford University Press.
- Hutto, D.D. 2004. The limits of spectatorial folk psychology. *Mind & Language* 19: 548–573.
- Hutto, D.D. 2008. *Folk psychological narratives: the sociocultural basis of understanding reasons*. Cambridge: MIT Press.
- Iacoboni, M., and M. Dapretto. 2006. The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience* 7: 942–951.
- Jacob, P. 2011. The direct-perception model of empathy: a critique. *Review of Philosophy and Psychology*. doi:10.1007/s13164-011-0065-0.
- Kobayashi, C., and E. Temple. 2009. Cultural effects on the neural basis of theory of mind. *Progress in Brain Research* 178: 213–223.
- Kovács, A., E. Teglas, and A. Endress. 2010. The social sense: susceptibility to others' beliefs in human infants and adults. *Science* 330: 1830–1834.
- Koyasu, M. 1997. The development of 'theory of mind' in young children: a comparison between mental representation and photographic representation. *Japanese Psychological Review* 40: 97–109.
- Lebra, T.S. 1993. Culture, self, and communication in Japan and the United States. In *Communication in Japan and the United States*, ed. W. Gudykunst, 51–87. Albany: State University of New York Press.
- Leslie, A.M. 1987. Pretense and representation: the origins of "theory of mind". *Psychological Review* 94: 412–426.
- Leslie, A.M. 2000. 'Theory of mind' as a mechanism of selective attention. In *The New Cognitive Neurosciences*, 2nd ed, ed. M. Gazzaniga, 1235–1247. Cambridge: MIT Press.
- Leslie, A.M., Y. German, and P. Polizzi. 2005. Belief-desire reasoning as a process of selection. *Cognitive Psychology* 50: 45–85.
- Lillard, A. 1998. Ethnopsychologies: cultural variations in theories of mind. *Psychological Bulletin* 123: 3–32.
- Lillard, A. 2002. Just through the looking glass: children's understanding of pretense. In *Pretending and Imagination in Animals and Children*, ed. R. Mitchell, 102–114. Cambridge: Cambridge University Press.
- Liu, D., H.M. Wellman, T. Tardiff, and M.A. Sabbagh. 2008. Theory of mind development in Chinese children: a meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology* 44: 523–531.
- Miller, J.G. 1984. Culture and the development of everyday social explanation. *Journal of Personality and Social Psychology* 46: 961–978.
- Mills, S. 2001. The idea of different folk psychologies. *International Journal of Philosophical Studies* 4: 501–519.

- Morris, M.W., and K. Peng. 1994. Culture and cause: american and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology* 67: 949–971.
- Mitchell, J.P. 2005. The false dichotomy between simulation and theory-theory: the argument's error. *Trends in Cognitive Science* 9(8): 363–364.
- Naito, M. 2003. The relationship between theory of mind and episodic memory: evidence for the development of autoeitic consciousness. *Journal of Experimental Child Psychology* 85: 312–336.
- Naito, M. 2007. Current status of theory of mind research and future directions. *Annual Review of Japanese Child Psychology* 47: 1–3.
- Naito, M., and K. Koyama. 2006. The development of false- belief understanding in Japanese children: delay and difference? *International Journal of Behavioral Development* 30: 290–304.
- Nichols, S., and S. Stich. 2003. *Mindreading. An integrated account of pretence, self-awareness, and understanding of other minds*. Oxford: Clarendon.
- Noë, A. 2004. *Action in Perception*. Cambridge, MA: MIT Press.
- Onishi, K.H., and R. Baillargeon. 2005. Do 15-month old infants understand false beliefs? *Science* 308: 255–258.
- Perner, J. 1991. *Understanding the representational mind*. Cambridge: Bradford Books/MIT-Press.
- Perner, J., and B. Lang. 1999. Development of theory of mind and executive control. *Trends in Cognitive Science* 3(9): 337–344.
- Perner, J., and T. Ruffman. 2005. Infants' insight into the mind: how deep? *Science* 308: 214–216.
- Perner, J., B. Lang, and D. Kloo. 2002. Theory of mind and self control: more than a common problem of inhibition. *Child Development* 73: 752–767.
- Rakoczy, H., M. Tomasello, and T. Striano. 2006. The role of experience and discourse in children's developing understanding of pretend play actions. *British Journal of Developmental Psychology* 24 (2): 305–335.
- Ratcliffe, M. 2007. *Rethinking commonsense psychology: a critique of folk psychology, theory of mind and simulation*. Basingstoke: Palgrave Macmillan.
- Reddy, V., and O. Morris. 2004. Participants don't need theories: knowing minds in engagement. *Theory and Psychology* 14: 647–665.
- Repacholi, B.M., and A. Gopnik. 1997. Early reasoning about desires: evidence from 14- to 18-month olds. *Developmental Psychology* 33: 12–21.
- Rietveld, D.W. 2008. *Unreflective action: a philosophical contribution to integrative neuroscience*, 2008. Amsterdam: ILLC Dissertation Series.
- Rosaldo, M.Z. 1980. *Knowledge and passion: Ilongot notions of self and social life*. Cambridge: Cambridge University Press.
- Ruffman, T., W. Garnham, A. Import, and D. Connolly. 2001. Does eye gaze indicate knowledge of false belief: charting transitions in knowledge. *Journal of Experimental Child Psychology* 80: 201–224.
- Saxe, R. 2005. Hybrid vigor: reply to Mitchell. *Trends in Cognitive Science* 9(8): 364.
- Scholl, B.J., and A.M. Leslie. 1999. Modularity, development and 'Theory of mind'. *Mind & Language* 14: 131–153.
- Scott, R.M., and R. Baillargeon. 2009. Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development* 80: 1172–1196.
- Song, H., and R. Baillargeon. 2008. Infants' reasoning about other's false perceptions. *Developmental Psychology* 44(6): 1789–1795.
- Southgate, V., A. Senju, and G. Csibra. 2007. Action anticipation through attribution of false belief by two-year-olds. *Psychological Science* 18: 587–592.
- Spaulding, S. 2010. Embodied cognition and mindreading. *Mind & Language* 25(1): 119–140.
- Straus, A. 1977. Northern Cheyenne ethnopsychology. *Ethos* 5: 326–352.
- Surian, L., S. Caldi, and D. Sperber. 2007. Attribution of beliefs to 13-month-old infants. *Psychological Science* 18: 580–586.
- Tooby, J., and L. Cosmides. 1995. Mapping the evolved functional organization of mind and brain. In *The Cognitive Neurosciences*, ed. M. Gazzaniga, 1185–1197. Cambridge: MIT Press.
- Vinden, P. 1996. Junin Quechua children's understanding of mind. *Child Development* 67: 1707–1716.
- Vinden, P. 1999. Children's understanding of mind and emotion: a multi-culture study. *Cognition & Emotion* 13: 19–48.
- Vinden, P. 2002. Understanding minds and evidence for belief: a study of Mofu children in Cameroon. *International Journal of Behavioral Development* 26(5): 445–452.
- Wellman, H.M. 1990. *The child's theory of mind*. Cambridge: MIT Press.
- Wellman, H.M. 1998. Culture, variation, and levels of analysis in folk psychologies. *Psychological Bulletin* 123: 33–36.

- Wellman, H.M. 2002. Understanding the psychological world: developing a theory of mind. In *Handbook of Childhood Cognitive Development*, ed. U. Goswami, 167–187. Oxford: Blackwell.
- Wellman, H.M., and A.T. Phillips. 2001. Developing intentional understandings. In *Intentions and Intentionality: Foundations of a Social Cognition*, ed. B.F. Malle, 125–148. Cambridge: MIT Press.
- Wierzbicka, A. 1992. *Semantics, culture, and cognition*. Oxford: Oxford University Press.
- Wierzbicka, A. 2009. Language and Metalanguage: key issues in emotion research. *Emotion Review* 1: 3–14.
- Wimmer, H., and J. Perner. 1983. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13: 103–128.
- Woodward, A.L. 1998. Infants selectively encode the goal object of an actor's reach. *Cognition* 69: 1–34.
- Woodward, A.L. 2005. The infant origins of intentional understanding. In *Advances in Child Development and Behavior Volume 33*, ed. R.V. Kail, 229–262. Oxford: Elsevier.