Solving the Puzzle about Early Belief-Ascription

KATHARINA A. HELMING, BRENT STRICKLAND AND PIERRE JACOB

Abstract: Developmental psychology currently faces a deep puzzle: most children before 4 years of age fail elicited-response false-belief tasks, but preverbal infants demonstrate spontaneous false-belief understanding. Two main strategies are available: cultural constructivism and early-belief understanding. The latter view (unlike the former) assumes that failure at elicited-response false-belief tasks need not reflect the inability to understand false beliefs. The burden of early-belief understanding is to explain why elicited-response false-belief tasks are so challenging for most children under 4 years of age. The goal of this article is to offer a pragmatic framework whose purpose is to discharge this burden.

Ever since the publication and discussion of Premack and Woodruff's (1978) seminal paper entitled 'Does the chimpanzee have a theory of mind?', false-belief understanding has been heralded as a distinctive hallmark of human social cognition.[1] But developmental psychology is presently confronted with a deep puzzle. On the one hand, two solid decades of experimental work have shown that not until they are at least 4 years old can the majority of young children successfully pass a variety of elicited-response false-belief tasks, in which they are being directly asked a question about an agent's false (or true) beliefs. On the other hand, more and more evidence from the past ten years or so shows that preverbal human infants spontaneously expect others to act in accordance with the contents of their true and false beliefs. The puzzle is: why do most children fail standard elicited-response false-belief tasks until they are at least 4 years old, while the looking behavior of preverbal infants strongly suggest that they can track the contents of others' false beliefs about an object's location?

As we shall argue in the first section, there are two broad answers to this puzzle. According to advocates of cultural constructivism, the human mindreading ability to

We are grateful to three anonymous reviewers for their comments on this paper, to Susan Carey, Peter Carruthers and Josef Perner for detailed comments on an earlier version and to Renée Baillargeon, Hanoch Ben-Yami, Ned Block, Clark Barrett, Steve Butterfill, Gergo Csibra, Fiery Cushman, Gyuri Gergely, Agi Kovacs, Yuyan Luo, Olivier Mascaro, Henrike Moll, Paul Pietroski, Victoria Southgate, Dan Sperber, Denis Tatone and Mike Tomasello for very useful conversations. The research reported in this manuscript was supported by the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013)/ERC grant agreement n° [609819], SOMICS.

Address for correspondence: Pierre Jacob, Institut Jean Nicod, École normale supérieure, Pavillon Jardin, 29, rue d'Ulm, 75005 Paris, France.
Email: jacobpiotr11@gmail.com

[1] Bennett (1978), Dennett (1978), Harman (1978) argued that false-belief understanding is a hallmark of theory of mind. So far, false-belief understanding seems to be distinctly human (cf. Call and Tomasello, 2008).

Mind & Language, Vol. 31, No. 4 September 2016, pp. 438–469.

track the contents of others' psychological states (in particular others' false beliefs) is a cultural skill that is transmitted from one generation to the next by means of verbal interactions.[2] Since advocates of cultural constructivism assume that only children who pass standard elicited-response false-belief tasks have the ability to track the contents of others' false beliefs, their burden is to offer low-level accounts of the infant findings consistent with their hypothetical inability to track the contents of others' false beliefs.

On the other hand, advocates of early-belief understanding (who are skeptical that older children learn to track the contents of others' false beliefs either on the basis of their own experience or from explicit teaching by competent adults) assume that the looking behavior and looking times of preverbal infants are evidence that they can track the contents of others' false beliefs. Their entirely different burden is to explain why elicited-response false-belief tasks are so challenging for most young children until they are 4 years old. We shall call it the 'early-belief understanding challenge'.

In the first section, we explain why we opt for the latter response to the puzzle about the discrepant developmental findings (based on early belief-understanding) and why failure at elicited-response false-belief tasks need not reflect failure to track the contents of others' false beliefs. The main goal of our article is to offer a pragmatic answer to the early-belief understanding challenge. In the second section, we spell out one of the prevalent responses to this challenge, i.e. the so-called 'processing-load' account, and argue that it is incomplete. In the third section, we articulate a pragmatic framework whose purpose is to more fully address the early-belief understanding challenge. Our framework rests on the following two distinctions: while humans can mindread the relevant psychological states of agents of both *instrumental* and *communicative* actions, they can further take either a *third*-person perspective or a *second*-person perspective on both kinds of actions. We argue that what makes some standard false-belief tasks particularly challenging for young children is that success on these tasks requires taking a *third*-person perspective onto the mistaken agent's instrumental action, while they must at the same time take a *second*-person perspective on the experimenter's communicative action and answer her direct question.[3]

1. Two Main Responses to the Discrepant Developmental Findings

Following the publication of Wimmer and Perner's (1983) well-known paper, much experimental work has been devoted to elicited-response change-of-location false-belief tasks.

[2] For further discussion, cf. Jacob, in press. On behalf of cultural constructivism, Heyes and Frith (2014) have recently proposed that human children learn to read others' minds the way they learn to read words. Cf. Strickland and Jacob (2015) for discussion.

[3] Cells 1 and 3 of Table 1.

For example, participants who know in which of two opaque containers a toy has last been placed are directly asked to predict where a mistaken agent will look for it (as in the Sally-Anne task). The evidence shows that most 3-year-olds point to the object's actual location, not to the empty location where the agent falsely believes the object to be.[4] This is known in the developmental literature as the 'reality bias' or the 'curse of knowledge' (cf. Birch and Bloom, 2003, 2004, 2007).

In contrast, more and more evidence based on so-called *spontaneous-response* tasks, in which participants are *not* requested to produce any answer in response to some direct question, suggests that preverbal human infants and older toddlers can track the content of an agent's false belief, as measured by such methods as violation-of-expectation and anticipatory looking. For example, in a seminal study, Onishi and Baillargeon (2005) demonstrated that 15-month-olds look reliably longer when an agent's action is incongruent rather than congruent with the content of her true or false belief. They looked longer when the agent reached for the empty location with a true rather than a false belief and when she reached for the toy's actual location with a false rather than a true belief. In an anticipatory looking paradigm, 25-month-olds were shown to first-gaze correctly towards the empty location where a mistaken agent believed her toy to be in anticipation of her action (Southgate *et al.*, 2007), thereby confirming the findings first reported by Clements and Perner (1994).

To solve the puzzle of these discrepant developmental findings, psychologists have offered two broad strategies, one of which assumes that failure at elicited-response false-belief tasks reflects the inability to ascribe false beliefs to others. On this view, the ability to ascribe false beliefs is taken to be the output of 'a cultural process tied to language acquisition' (Perner and Ruffman, 2005, p. 214). The burden of this cultural constructivist strategy is to explain away the findings about preverbal infants without crediting them with the ability to ascribe a false belief to an agent. Other psychologists (including Baillargeon *et al.*, 2010; Leslie, 2005; Leslie *et al.*, 2004; Leslie *et al.*, 2005; Scott *et al.*, 2010), who support early-belief understanding, take findings about preverbal infants at face value as showing that they can ascribe false beliefs to others. Their burden is to explain why elicited-response false-belief tasks are so challenging for 3-year-olds. The prevalent non-constructivist explanation is the processing-load account first anticipated to some extent by Fodor (1992) by Leslie (1994), Leslie *et al.* (2004 and 2005), and recently advocated by Baillargeon and colleagues.[5]

---

[4] This version of the elicited-response false-belief task known as the Sally-Anne task was adapted by Baron-Cohen *et al.* (1985) from Wimmer and Perner's (1983) design. At the end of Section 4, we shall examine two other kinds of elicited-response false-belief tasks: on the one hand, there are change-of-location false-belief tasks in which participants do not know an object's location and must determine it from the testimony of a mistaken agent. On the other hand, there are so-called 'unexpected-contents' false-belief tasks.

[5] In so-called 'change-of-location' false-belief tasks, the mistaken agent is performing an instrumental action whose goal is to retrieve some object. So the content of the agent's false belief is

In the next few paragraphs of this section, we will briefly explain why we think that early-belief understanding is a more promising resolution of the puzzle of the discrepant developmental findings than any current version of cultural constructivism.

The burden of cultural constructivism is to explain away a wide range of data based on spontaneous-response tasks congruent with early-belief understanding. Ten years before the publication of the seminal paper by Onishi and Baillargeon (2005), Clements and Perner (1994) crucially reported that most 3-year-olds who incorrectly pointed to the toy's actual location when asked to predict where the mistaken agent will nevertheless look for her toy accurately gazed at the empty location. In the ten years following Onishi and Baillargeon's (2005) paper, more and more evidence based on spontaneous-response tasks has been reported with younger and younger children that is consistent with early-belief understanding (Surian *et al.*, 2007; Kovacs *et al.*, 2010; Luo, 2011; Southgate and Vernetti, 2014).[6]

There are presently three purported low-level accounts, all of which share the cultural constructivist assumption that failure at elicited-response false-belief tasks is best interpreted as directly reflecting the inability to track the contents of others' false beliefs. According to the straight associationist account—first sketched by Perner and Ruffman (2005) and recently updated by Heyes (2014) into a low-level perceptual novelty account—in change-of-location false-belief tasks, young children are supposed to form three-way associations between an agent, a toy and a location, and to expect these associations to persist in space and time. The teleological account (advocated by Perner and Roessler, 2010, Perner and Roessler, 2012, and Roessler and Perner, 2013) posits that young children reason about others' actions according to the teleological principles of fact-based objective rationality, not according to the

about the object's location. In elicited-response tasks, participants' understanding of the contents of others' false beliefs is tested by means of their (verbal or non-verbal) answer to a direct question asked by the experimenter. In spontaneous-response tasks, participants' understanding is inferred from their spontaneous behavior (e.g. looking time or anticipatory gaze) in response to the test trial that can be either congruent or incongruent with the familiarization trials (cf. Baillargeon *et al.*, 2010; He *et al.*, 2012; Scott and Baillargeon, 2009; Scott *et al.*, 2010). Some tasks in which participants are invited to help a mistaken agent find her toy do not exactly fit the elicited-response/spontaneous-response distinction (cf. Buttelmann *et al.*, 2009 and Southgate *et al.*, 2010). Our two-by-two pragmatic framework explains why the above binary methodological distinction is both necessary but not sufficient to account for all the findings about young children and preverbal infants. In a nutshell, what is crucial to explaining the discrepant developmental findings is whether and how an infant's ability to track the content of another's false belief is being disrupted by the specific demands made by the experimenter's communicative action (see the beginning of Section 3 and Table 1).

[6] While Luo (2011) tested 10-month-olds, Kovacs *et al.* (2010) tested 7-month-olds, and Southgate and Vernetti (2014) tested 6-month-olds. Furthermore, the evidence about early-belief understanding is not restricted to understanding the contents of others' false beliefs about an object's location, but extends to the contents of others' false beliefs about unexpected-contents (He *et al.*, 2011; Buttelmann *et al.*, 2014), non-obvious properties (Scott *et al.*, 2010) and object-identity (Scott and Baillargeon, 2009).

mentalistic principles of subjective rationality (which would include the computation of the content of an agent's false belief). For example, if an agent's goal is to find a toy and if children are teleologists (in Perner and Roessler's sense), then children will predict that the agent will look for the toy at its actual location (an objective fact and thus an objective reason), despite the agent's false belief about its location (a subjective reason). Finally, according to the two-systems model of mindreading (advocated by Apperly, 2011; Apperly and Butterfill, 2009; Butterfill and Apperly, 2013a), humans make use of two systems for mindreading others' psychological states: an early developing system available to very young children and a later-developing system unavailable to them. While the latter enables older children and adults to track the contents of others' true and false beliefs, the former enables young children to track the contents of others' true and false registrations, which are taken to be ontogenetic (and perhaps phylogenetic) precursors of true and false beliefs.[7]

All three versions of cultural constructivism face at least two main kinds of objections. First, they all rest on the fundamental assumption that failure at elicited-response false-belief tasks reflects the inability to ascribe false beliefs to others. Now, failure at elicited-response false-belief tasks could directly reflect the inability to ascribe false beliefs to others only if this ability were in turn sufficient for success at elicited-response false-belief tasks. But as several critics of cultural constructivism have pointed out (Bloom and German, 2000; Leslie *et al.*, 2004; Leslie *et al.*, 2005; Scholl and Leslie, 1999), it is quite clear that being able to track the contents of others' false beliefs is far from sufficient for success at elicited-response change-of-location false-belief tasks. First of all, one must understand the language spoken by the experimenter who asks the question.[8] Secondly, one must have the pragmatic ability to understand that what the experimenter is asking is where Sally *will* look for her toy, not e.g. where she *should* look for it. Furthermore, one must have enough executive resources in order to inhibit the potential propensity to answer the experimenter's question on the basis of one's own knowledge of the toy's location (for recent meta-analytic investigation of the role of executive functions in false-belief tasks cf. Devine and Hughes, 2014). So, there are strong reasons to doubt that failure at elicited-response change-of-location false-belief tasks should be taken to directly reflect the inability to track the contents of others' false beliefs about an object's location, in accordance with cultural constructivism.

---

[7] Butterfill and Apperly (2013a) argue that infants and adults under cognitive load track others' beliefs about an object's location, not by representing beliefs as such, but instead by tracking others' registrations. In general, we agree that it is possible to track one property (e.g. toxicity), not by representing it as such, but instead by tracking another correlated property (e.g. odor). However, since we do not accept Butterfill and Apperly's application of this distinction to the case of beliefs, we freely talk of tracking (or representing) the contents of others' beliefs without committing ourselves to any contrast between tracking beliefs and representing them as such.

[8] If the question is asked in English, then a monolingual Russian speaker with the full ability to track the contents of others' false beliefs can only fail the elicited-response change-of-location false-belief task. For further recent meta-analytic investigation of the role of language understanding in false-belief tasks, cf. Milligan *et al.*, 2007.

Secondly, it is no trivial task for any of the current versions of cultural constructivism to meet the main challenge of offering low-level accounts of the infant findings consistent with their hypothetical inability to track the contents of others' false beliefs. To see why, consider a recent study by Senju *et al.* (2011) based on the anticipatory looking paradigm and first suggested by Heyes (1998) and exploited by Meltzoff and Brooks (2008). In the familiarization stage, 18-month-olds experience the effect of wearing either an opaque blindfold through which they cannot see or a trick blindfold through which they can see. But the difference between an opaque and a trick blindfold is undetectable for an outside observer. In the first trials of the test phase, the children are familiarized to seeing an agent retrieve her toy at the location where a puppet has placed it in front of her. The agent's action is always preceded by a pair of visual and auditory cues. In the last test trial, the agent sees the puppet place the toy in one of the two boxes; she ostensibly covers her eyes with a blindfold, and the puppet removes the toy. After the puppet disappears, the agent removes her blindfold and the cues are produced. Using an eye-tracker, Senju *et al.* (2011) found that only infants who had experienced an opaque blindfold, not infants who had experienced a trick see-through blindfold, reliably made their first saccade towards the empty location in anticipation of the agent's action.

As Senju *et al.* (2011, p. 879) argue, their findings are hard to reconcile with the associationist account: since the opaque and the trick blindfolds could not be distinguished from the perspective of an outside observer, all infants saw exactly the same events. Since they saw the same events, they should have formed exactly the same threefold association between the agent, the toy and the location, and on this basis they should have gazed at the same location in anticipation of the agent's action. But they did not. Infants whose view had been previously obstructed by an opaque blindfold, not infants whose view had not been obstructed by a trick blindfold, expected the blindfolded agent to mistakenly believe that the object was still in the opaque container after the puppet removed it.[9]

These findings also seem hard to reconcile with the claim by Perner and Roessler (2010) that infants are teleologists and can reason about others' actions only in accordance with the fact-based principles of objective rationality, not in accordance with the principles of subjective rationality. Although there is an objective difference

[9] Heyes (2014, p. 657) entertains the alternative non-mentalistic possibility that children in the opaque blindfold condition were more distracted than children in the trick blindfold condition when they saw the agent wear a blindfold in the test trial. As a result, they failed to notice that the toy had been removed and gazed at the location where they (not the agent) falsely thought the toy still was. Given that the visual appearances of the opaque and the trick blindfolds were indistinguishable by external observers, she must assume that wearing opaque blindfolds rather than trick blindfolds must have disposed children to be selectively distracted when seeing an agent wear a blindfold. Heyes is thereby calling for further controls. However, unless she provides an explanation or a justification for why seeing the agent wear a blindfold in the test trial should selectively impair the ability of children who experienced opaque blindfolds, but not trick blindfolds, in the familiarization trial, to attend to subsequent events in the test trial, her interpretation of the findings seems rather ad hoc.

between the fact that in one condition the agent is wearing an opaque blindfold and in the other condition the agent is wearing a trick blindfold, this difference cannot be ascertained by an outside observer. Infants must have formed different expectations about the agent's likely action on the basis of representing the content of the agent's belief about the object's location. They must in turn have formed expectations about the agent's belief on the basis of their own earlier subjective visual experience arising from wearing either an opaque or a trick blindfold, not on the basis of objective fact-based reasons.

These experimental findings also raise a dilemma for the two-systems model of mindreading, which rests on the fundamental distinction between tracking an agent's *registration* and representing an agent's *belief as such*. Infants and adults under cognitive load are credited with the former, not with the latter, ability. The question is: could an agent's registration be really distinct from an agent's belief about an object's location and still count as a genuine epistemic psychological state? Does tracking another's registration rest on genuine 'minimal' mindreading or is it part of behavior reading?[10]

An agent is officially said to stand in the registering relation to an object and a location only if the agent *encountered* the object at that location. An agent is further said to stand in the encountering relation to an object if the object stood in the agent's field and was not occluded from the agent's line of sight (Apperly and Butterfill, 2009, p. 962). Advocates of the two-systems model must make a choice: either they assume that registration is a purely extensional non-psychological relation between an agent, an object and a location, or they do not. Presumably, registration could only be an extensional relation if encountering was and if registration inherited its extensionality from encountering. Now, only if the encountering relation fully rests on the existence of an unobstructed geometrical straight line between the eyes of the relevant agent and the relevant object can the encountering relation, and thus the registration relation, count as extensional non-psychological relations. In some of their writings, the advocates of the two-systems model are clearly tempted by the first horn of the dilemma.[11] But if tracking another's registration just results from checking whether the line of sight between the agent's eyes and the toy is obstructed or not, then tracking another's registration will fail to explain why infants in the opaque blindfold condition and infants in the trick blindfold condition had different expectations about the agent's likely action. Only by taking into account the non-extensional difference between infants having *experienced* either an opaque

[10] Cf. Heyes's (1998) seminal paper and the papers by Spaulding (2013) and Zawizki (2013) and the responses by Butterfill and Apperly (2013b).

[11] As Apperly and Butterfill (2009, p. 962) put it, 'the key requirement is that conditions under which an encounter occurs must be specified without appeal to anything psychological'. As Butterfill and Apperly (2013a, p. 616) further put it, 'since encounterings are relations not representations (by definition), representing encounterings will differ from representing perceptions in that only the latter involves representing representations'.

or a trick blindfold could one explain their different expectations about the agent's action. So far then, given the findings reported by Senju *et al.* (2011), the burden of proof clearly rests on the advocates of cultural constructivism.

2. The Scope and Limits of the Processing-Load Account

The alternative approach to the puzzle of the discrepant developmental findings interprets infants' data as evidence that they can track the contents of others' false beliefs. The burden of this alternative approach is to explain why elicited-response false-belief tasks are so challenging for most children until they are 4 years old (the early-belief understanding challenge). One of the most explicit answers to this challenge is the processing-load account (Baillargeon *et al.* 2010; Scott and Baillargeon, 2009; Scott *et al.*, 2010; He *et al.*, 2012; Baillargeon *et al.*, 2015).[12] According to the processing-load account, success at elicited-response change-of-location false-belief tasks involves at least three processes, the first of which is the representation of the content of the agent's false belief. The second process is a response-selection process whereby participants must access the content of the agent's false belief. The third process is a response-inhibition process whereby participants must inhibit any prepotent tendency to answer the test question based on their own knowledge. Thus, participants must have enough executive resources in order to inhibit any such propensity. As Baillargeon *et al.* (2010, p. 115) have put it, while spontaneous-response tasks involve only the false-belief-representation process, 'young children fail elicited-response tasks because simultaneously executing the false-belief-representation, response-selection, and response-inhibition processes overwhelms their limited resources' (for detailed discussion, cf. Carruthers, 2013). In short, failure in an elicited-response false-belief task need not reflect failure to represent the contents of others' false beliefs.

On our view, the main problem with the processing-load account is not that it is wrong, but that so far it is incomplete, as a recent finding based on a novel false-belief task (cf. Rubio-Fernández and Geurts, 2013) illustrates. It involves a puppet (the Duplo girl) who has a false belief about the location of her bananas. 3-year-olds, who know the actual location of the bananas, are prompted to act out the puppet's most likely action by the experimenter who tells them: 'What happens next? You can take the girl yourself if you want. What is she going to do now?' In response, the majority of 3-year-olds move the girl to the empty location in accordance with the content of her false belief.

The question is: why did not most 3-year-olds take the puppet to her bananas? Why did the experimenter's prompt not overwhelm their inhibitory resources and

[12] For an early anticipation of the basic insights of the processing-load account that predates the experimental findings of Onishi and Baillargeon (2005), cf. Leslie (1994), Scholl and Leslie (2001).

generate a reality bias, as predicted by the processing-load account? In order to take the Duplo girl to the empty location, 3-year-olds must have inhibited their own knowledge and selected their representation of the content of the girl's false belief. Clearly, the processing-load account needs some further explanation of why being asked the where-prediction question, but not being prompted to act out the mistaken agent's next action, overwhelms young children's inhibitory resources and generates a reality bias. As we shall shortly argue, *pragmatic* factors are likely to play a major role in this further explanation.

While Scott *et al.* (2012) explicitly endorse the processing-load account, they also tentatively further sketch a pragmatic approach to the failure of young children in elicited-response false-belief tasks: 'when children are asked the test question (and thus shift from merely observing the test scene to participating in a conversation about it), their own perspective on the scene naturally becomes prominent and must be inhibited to allow them to adopt the agent's perspective' (p. 190). They thereby seem to recognize that a pragmatic account may be needed to explain why only young children's true belief about the object's actual location, not their true belief about the content of the mistaken agent's false belief about the object's location, makes its way into their answer to the where-prediction question. Recently, Carruthers (2013, p. 153) has offered what he calls a 'triple-load' account, according to which success at elicited-response false-belief tasks requires processing the speech of the experimenter, figuring out her underlying communicative intention, and further formulating an action that would serve to fulfill the communicative agent's informative intention.

Siegal and Beattie (1991) were the first to hypothesize and test a pragmatic approach. Their conjecture was that in the pragmatic context in which the where-prediction question is being asked, the experimenter's intended question is about the location at which the mistaken agent will *first* look for the object. They further hypothesized that the temporal meaning of the experimenter's question is conveyed by a Gricean implicature that young children fail to retrieve. They found that while only 35% of 3- and 4-year-olds correctly answered the where prediction question, 71% did when the temporal meaning of the experimenter's utterance was linguistically encoded by the use of the temporal adverb 'first' ('Where will Sally look for her marble first?'). While we are not convinced that the where-prediction question is meant to ask where the mistaken agent will look for the object *first* nor that this temporal aspect of the speaker's meaning is implicitly conveyed by means of a Gricean implicature, we fully agree that 3-year-olds' failure in elicited-response false-belief tasks cries for a detailed pragmatic framework.

3. The Present Framework

Our pragmatic framework rests on two fundamental distinctions. On the one hand, humans are able to make sense of two kinds of agency: instrumental and ostensive communicative agency. On the other hand, we argue that humans can

take two distinct perspectives on another's action: a third-person perspective and a second-person perspective.[13]

Understanding an agent's instrumental action requires tracking the contents of her motivation (e.g. her desire) and her epistemic state (e.g. her belief), and construing her action as an efficient means of fulfilling her motivation in light of her belief about the local constraints imposed to the agent by her immediate environment (cf. Gergely and Jacob, 2013). As argued by Sperber and Wilson (1986) and Wilson and Sperber (2004), who simplified Grice's earlier (1969) account, an agent who performs an ostensive communicative action has two complementary intentions both of which are directed to a specific recipient. First, she has some *informative* intention to cause her recipient (or addressee) to acquire a new belief or a new desire. She may intend to make manifest either some fact that her recipient may come to believe or some possible state of affairs that her recipient may come to wish to turn into an actual state of affairs by his action. (What makes a communicative action *ostensive* is that the communicative agent produces an ostensive signal—e.g. direct gaze or speech—whose purpose is to provide her recipient with evidence that she has a communicative intention.) Secondly, she has the higher-order *communicative* intention to make her *informative* intention manifest to her recipient. The agent's informative intention is fulfilled if her authority is sufficient to cause the recipient to take the state of affairs made manifest by her communicative act either as a fact which he believes to obtain or as a desirable goal for his own action. The agent's informative intention is thwarted otherwise.

Ostensive communicative actions are the hallmarks of human cooperative interactions. The agent has a particular recipient in mind and makes manifest her intention to change one of his mental states on the assumption that the recipient's mind is open to such a change. The recipient's task is to fulfill the agent's goal as a function of the reliability and/or authority of the agent. Since the goal of the agent's communicative act is to change a particular recipient's mind, the full success of the agent's communicative act requires the recipient to take a *second-person perspective* on the agent's communicative act by sharing responsibility for the success of the agent's goal, which is that the recipient fulfill the agent's informative intention. In short, the recipient is requested not only to fulfill the agent's communicative intention by recognizing her informative intention, but also to fulfill her informative intention by either accepting a new belief or a new desire. While the recipient takes a second-person perspective

[13] In the past few years, so much work has been devoted to the second-person perspective, ranging from the role of second-person reasons in moral philosophy (e.g. Darwall, 2006) to the role of primary dyadic interactions in human social cognition (e.g. Schilbach *et al.*, 2013), that Naomi Eilan published a (2014) paper entitled 'The You Turn'. Here, we offer our own account of the distinction between taking a second- and a third-person perspective on another's action, which can either be an instrumental or an ostensive communicative action. We believe our account to be more precise than other current attempts and especially more relevant to clarifying the developmental puzzle about early belief-ascription. In particular, it enables us to explain why the methodological distinction between elicited-response and spontaneous-response false-belief tasks is necessary but not sufficient to explain the discrepant developmental findings.

on the agent's communicative act, it is also open to him to decline to fulfill the agent's informative intention and refrain from endorsing a new belief or a new desire.

Thus, to take a *second-person* perspective on the agent's ostensive communicative act is to recognize that the success or the failure of the agent's act also depends on one's own response. By contrast, a by-stander who is not the recipient can take a detached *third-person perspective* on an agent's communicative act and recognize the agent's informative intention that her recipient acquires a new belief or a new desire *without* feeling responsible for the success or the failure of the agent's communicative act by fulfilling or not the agent's informative intention.

One can similarly take either a third- or a second-person perspective on an agent's instrumental action.[14] To make sense of an agent's instrumental action from a detached *third-person perspective* is to track the content of her motivation and epistemic state *without* intending to contribute to either the success of the failure of the agent's instrumental action. (Spontaneous-response change-of-location false-belief tasks typically require participants to take a detached third-person perspective onto a mistaken agent's instrumental action.) But one can also take a *second-person perspective* onto an agent's instrumental action if one assumes that the success or the failure of the agent's action depends also on one's own contribution. Thus, one can take a second-person perspective onto an agent's instrumental action by either *cooperating* or *competing* with the agent. One cooperates with an agent if one intends to contribute to the success of the agent's goal by *helping* the agent achieve her goal. One competes with an agent if one intends to contribute to the failure of the agent's goal by *hindering* the agent from achieving her goal.

The goal of our pragmatic approach is to fill the remaining gaps in the answer provided by the processing-load account to the question: what makes elicited-response change-of-location false-belief tasks so challenging for most children under 4 years of age? Our pragmatic approach rests on the four distinct possibilities afforded by the human mindreading system. On the one hand, humans can mind read the psychological states of agents of both instrumental and communicative actions. On the other hand, they can take either a detached third-person perspective or an engaged second-person perspective on either an instrumental or a communicative action performed by another individual. We summarize the four distinct possibilities in Table 1.

	Third-person	**Second-person**
Instrumental action	**1**	**2**
Communicative action	**3**	**4**

Table 1 *Four possibilities afforded by the human mindreading system.*

[14] In many experimental papers (e.g. Moll and Tomasello, 2007; Moll, Carpenter and Tomasello, 2007) and in his (2014) book, Mike Tomasello has fruitfully exploited the distinction between taking a third- and a second-person perspective on another's action.

As we shall argue, the systematic failure of most children under 4 years of age in elicited-response change-of-location false-belief tasks reflects, not their inability to track the contents of others' false beliefs, but rather their failure to stick to a *third-person perspective* on the *instrumental* action performed by a mistaken agent while being requested to take a *second-person perspective* onto the experimenter's *communicative* action.[15]

In the rest of this section, we shall first review the evidence revealing the full scope of the ability of young children to take a detached third-person perspective onto others' actions (cells 1 and 3 of Table 1). Then we shall review the evidence showing that very young human children can take a second-person perspective onto another's nonverbal ostensive *communicative* action (cell 4 of Table 1). Finally, we shall review the evidence showing that they can also take a second-person perspective onto another's *instrumental* action and be spontaneously motivated to help (cell 2 of Table 1).

3.1 The Full Scope of Early Third-Person Perspective on Others' Actions

There is more and more evidence based on spontaneous-response tasks showing that very young infants form correct expectations about the true and false beliefs of agents of *instrumental* actions, from a detached and undisturbed *third-person perspective* on the agent's action (cell 1 of Table 1) (Onishi and Baillargeon, 2005; Southgate *et al.*, 2007; Luo, 2011).

Furthermore, a recent study by Scott *et al.* (2012) suggests that from a *third-person perspective* young children are even capable of making sense of the experimenter's *communicative* action. In a violation-of-expectation paradigm, 2.5-year-olds watched while an adult subject was submitted to an elicited-response change-of-location false-belief task enacted by two experimenters, one of whom was the mistaken agent called Sally, and the other of whom was the interrogating experimenter. In the familiarization trial, the children saw Sally play with the toy, place it in a green box located next to a blue box, in front of the experimenter and the subject, and leave. In the test trial, while Sally was away, the experimenter moved the toy from the green to the blue box and then asked the subject in the middle where Sally would look for her toy on her return. In response, the subject pointed either to the blue box (the toy's

[15] By drawing the distinction between taking a second-person and a third-person perspective on another's instrumental action, we adopt a framework distinct from Rubio-Fernández and Geurts's (2013, p. 28), for whom tracking another person's perspective 'is merely to say that [one] can form expectations about that person's actions based on observations of his or her behavior'. We do not fully agree that 'it is immaterial whether this capacity [to track another's perspective] involves … representations of beliefs'. For us, taking a third-person perspective on another's instrumental action requires tracking the contents of both the agent's relevant epistemic state (or belief) and her relevant motivation (or desire). Furthermore, our two-by-two pragmatic framework also explains why, unlike the standard-elicited response Sally-Anne task, the Duplo task does *not* generate what we call the 'referential bias'.

actual location) or to the green box (the box in which Sally falsely believed her toy to be). Scott *et al.* (2012) report that 2.5-year-olds looked reliably longer when the subject pointed to the toy's actual location than to the empty location.

This finding based on a spontaneous-response task shows how extended the scope of children's third-personal spectatorial psychology is.[16] While 2.5-year-olds are known to fail elicited-response change-of-location false-belief tasks, they are able not only to track the content of the instrumental agent's (Sally's) false belief about the location of a toy, but also to form correct expectations about the interrogated subject's expectation about the content of Sally's false belief about the toy, from a purely detached third-person perspective. They are even capable of making sense of the experimenter's communicative action addressed to the interrogated subject, from a detached third-person perspective (cells 1 and 3 of Table 1).

3.2 Early Second-Person Perspective Onto Communicative Agency

There is also growing evidence that very young children are tuned to taking a second-perspective on others' ostensive communicative actions (cell 4 of Table 1). An *ostensive* stimulus is a signal produced by a communicative agent to provide her recipient with evidence that she has some communicative intention. One immediate effect of the presence of an ostensive signal is that the recipient knows that he is being directly addressed by a communicative agent. Much evidence shows that right after birth infants display selective sensitivity to such nonverbal ostensive stimuli as eye-contact, infant-directed speech (or motherese), and infant-directed contingent reactivity to their own responses (cf. Csibra, 2010; Csibra and Gergely, 2009; Gergely and Jacob, 2013 for reviews). The detection of ostensive signals generates referential expectations in infants. For example, 6-month-olds follow an agent's gaze shift to one of two objects, only if it is preceded by eye contact or infant-directed speech addressed to the infant (Senju and Csibra, 2008). After being cued by eye-contact, 12- and 8-month-olds expect to find an object at a location at which the communicative agent shifted her gaze and they look longer if the location to which the agent shifted her gaze turns out to be empty rather than non-empty (Csibra and Volein, 2008).

Csibra and Gergely (2009) have further argued that the reception of ostensive signals directed towards preverbal infants may cue them towards assuming a particular kind of second-person perspective onto a care-taker's communicative action, which they call the *pedagogical stance*, whereby they are disposed to interpret the nonverbal action of a communicative agent as conveying generalizable (or generic) information rather than episodic information. For instance, in a classical hide-and-seek game (first investigated by Piaget, 1954), an adult repeatedly hides a toy under container A until she changes and hides it under B, in plain view of 8-month-olds, whose task

[16] 'Spectatorial psychology' is the expression disparagingly used by Hutto (2004).

is to retrieve the toy. Young children are known to perseverate and keep looking under A. Topal *et al.* (2008) have shown that this perseveration error significantly decreases when children receive no ostensive signals from the adult. Arguably, when ostensively cued, young children do not interpret the situation as an episodic game in which they must retrieve the toy that has been hidden by the adult at some temporary location, but as a teaching session during which they are expected to learn some generalizable information about the toy's proper location.

Such evidence for natural pedagogy shows that preverbal infants are uniquely sensitive to others' communicative intentions, that they can spontaneously take an appropriate second-person perspective on an agent's ostensive communicative action and even fulfill her informative intention to cause them to acquire a new belief about some generic information.

3.3 Early Second-Person Perspective on Instrumental Agency

Spontaneous-response tasks probe the ability of young children to make sense of the motivations and epistemic states of agents of instrumental actions, from a detached third-person perspective (cell 1 of Table 1). Humans can also take a second-person perspective on another's instrumental action: they can cooperate with others and help them achieve their goals or else compete with them and prevent them from achieving their goals (cell 2 of Table 1).[17]

Much evidence shows that young children are altruistically inclined to take a second-person perspective on another's instrumental action and help the agent achieve her goal. As Tomasello (2009, pp. 5–6) has argued, the possibilities of altruistic helping vary along two different dimensions. On the one hand, helping another by sharing food, by fetching an out-of-reach object and by sharing information have different motivational costs for the helper. On the other hand, there are two basic cases in which one can help an agent involved in the execution of an instrumental action, according to whether the agent is known either to be trying and failing to achieve her goal for lack of skill or to have a deficient epistemic state (a state of ignorance or a false belief). While Warnecken and Tomasello (2006, 2007, 2009) have shown that young children are inclined to help an agent whom they see trying and failing to achieve a goal, Liszkowski *et al.* (2006) offer evidence that 18- and even 12-month-olds are motivated to point to objects which they believe an adult is looking for. More recently, Knudsen and Liszkowski (2012) have shown that 24- and 18-month-olds spontaneously point to an object's location for the benefit of an agent, if but only if the agent's goal is to retrieve the object and the

[17] Hamlin, Wynn and Bloom (2007) provide evidence that 6- and 10-month-olds prefer an agent who helps over one who hinders another achieve her goal. Notice that taking a second-person perspective on an agent's instrumental action in our sense is not limited to helping, but extends to hindering, the agent's fulfillment of her goal.

agent has a false belief about its location, not if the agent either knows the object's location or her goal is not to retrieve it (cell 2 of Table 1).

In two further experiments, young children were requested to take a second-person perspective on the *instrumental* action of an agent who holds a false belief about an object's location. First, Buttelmann *et al.* (2009) found that, when prompted to help an agent who was unsuccessfully trying to open a pink box which she mistakenly believed to contain her toy, 25- and even 18-month-olds, who knew that the toy was really in a yellow box, reliably approached the yellow box, not the pink box (cell 2 of Table 1). Secondly, Southgate *et al.* (2010) found that when 17-month-olds are requested to give one of two objects that is referred to as 'a sefo' by an agent who has a false belief about its location, and who points to the wrong location, they give the agent the intended referent of her pointing gesture, not the object at the demonstrated location (cell 2 of Table 1). These results show that before the end of their second year, children can track the content of an agent's false belief when this is needed for successfully helping an agent achieve the goal of her instrumental action from a second-person perspective.

Finally, in a standard elicited-response change-of-location false-belief task conducted by Matsui and Miura (2008), 3-year-olds saw a rabbit facing an apple and two colored boxes: a yellow and a blue box. The rabbit placed the apple into the yellow box and left. After seeing a panda come in, move the apple from the yellow into the blue box and leave the scene, children were asked the question: 'Where will the rabbit look for the apple?' The majority of 3-year-olds incorrectly pointed to the blue box (i.e. the apple's actual location). In a second elicited-response 'helping-false-belief' task, 3-year-olds saw the rabbit place an apple and an orange in the yellow box and leave. After seeing the panda come in, move the fruits into the blue box, both puppets reappeared and the children were asked the question: 'Who should we help find the fruits?' A majority of 3-year-olds correctly pointed to the rabbit. In a third elicited-response 'deceiving-false-belief' task built on the same scenario as the 'helping-false-belief' task, children were asked: 'Who should we deceive to keep the fruits?' 3-year-olds were at chance: half correctly pointed> to the panda.

To either help or deceive an agent involved in executing some instrumental action is to take a second-person perspective on the agent's action (cell 2 of Table 1). The last pair of findings suggests that it is cognitively less demanding to help a mistaken agent than to deceive a non-mistaken one. Arguably it is less demanding to point to an object's actual location (in accordance with the content of one's own knowledge) than to the empty location (in contradiction with the content of one's own knowledge). If young children can track the content of another's false belief about a toy's location, then they are more likely to efficiently help a mistaken agent find her toy than to predict her most likely action. They can efficiently help a mistaken agent by pointing to the toy's actual location in accordance with the content of their own knowledge (cf. Buttelmann *et al.*, 2009, and Southgate *et al.*, 2010). But if most point to the toy's actual location in accordance with the content of their own knowledge, then they will fail to accurately predict where a mistaken agent

will look for her toy (cf. Wimmer and Perner, 1983; Baron-Cohen *et al.*, 1985; Wellman *et al.*, 2001).

As it turns out, some evidence also suggests that being actively involved in causing another to have a false belief improves young children's performance in standard elicited-response change-of-location false-belief tasks (as recognized by Wellman *et al.*, 2001). For example, children's ability to act out the Duplo girl's action in accordance with the content of her false belief may be enhanced by their being enrolled into deceiving the girl by the experimenter whose action is described by Rubio-Fernández and Geurts (2013) as 'secretive'.[18] Moreover, a majority of young children have recently been shown to pass successfully elicited-response false-belief tasks when they also become able to correctly reason from others' false testimony (Mascaro and Sperber, 2009). Thus, children's developing ability to deceive and to deal with others' false testimonies may contribute to their success at elicited-response false-belief tasks.

4. Towards a Pragmatic Solution to the Early-Belief Understanding Challenge

The core of our pragmatic account is that young children are required to take a *second-person perspective* on the experimenter's *communicative* action, while they are simultaneously required to stick to a *third-person perspective* on the mistaken agent's *instrumental* action. This tension between the second- and the third-person perspectives generates two *biases*, one of which highlights the *epistemic* perspective on the object's actual location that is shared by the children and the experimenter, at the expense of the mistaken agent's different epistemic perspective on the object's location. The other bias *motivates* young children to help the mistaken agent achieve the goal of her instrumental action.

Before spelling out these biases, we review a couple of recent experiments that show how young children's spontaneous ability to track a mistaken agent's epistemic perspective on an object's location can be disrupted by being the addressees of the experimenter's where-prediction question. One shows that young children's ability to correctly anticipate where a mistaken agent will look for the object depends on whether or not they are *directly* addressed by the experimenter's utterance of the where-prediction question. The other shows that a majority of adults first look at the actual, not the empty location, before correctly answering the where-prediction question. Finally, in this last section, we shall demonstrate how our pragmatic account can accommodate these findings and can be extended to other variants of elicited-response false-belief tasks, which the majority of 3-year-olds have been also shown to fail.

[18] As Rubio-Fernandez and Geurts (2013, p. 30) put it: 'to the extent that deception facilitates false-belief reasoning … , it does so because deception helps children stay tuned to the perspective of the character who is being deceived'.

4.1 Disrupting Early Third-Person Perspective on Instrumental Agency

He *et al.* (2012) have recently combined the anticipatory looking paradigm with a verbal prompt. A first experimenter E1 was caused to have a false belief about her scissors' location (after she had placed them in one of two containers) by a second experimenter E2, who removed the scissors and placed them into her own pocket, in the absence of E1. In the false-belief test condition, E2 looked at the ceiling, chin in hand and uttered the sentence as if thinking out loud: 'But when E1 comes back, she's going to need her scissors again'. After a 2 s pause, she said: 'Where will she think they are?' In the knowledge condition, E1 knew where her scissors were because E1 was present when E2 switched the location of her scissors. In the question-false-belief condition, everything was the same as in the false-belief condition except that when E2 finally uttered the very same couple of sentences, she kept looking at the children's eyes as if addressing them.

He *et al.* (2012) coded the direction and measured the duration of children's gaze on the empty container, the non-empty container and E2 immediately before and after the verbal prompt. They found that 2.5-year-olds looked reliably longer at the empty container in the false-belief condition than in both the knowledge and the question-false-belief condition. (In other words, they correctly gazed at the empty container in the situation exemplified by cells 1 and 3 of Table 1, not in the situation exemplified by cells 1 and 4 of Table 1.) This finding confirms that 2.5-year-olds are able to gaze at the empty location where an agent falsely believes her target to be. It also suggests that this ability is being disrupted by a speaker's verbal prompt if young children take the speaker to request an answer from them. In other words, whether young children take a second- or a third-person perspective on the speaker's utterance of one and the same sentence makes all the difference to their ability to keep track of the content of the mistaken agent's false belief about the object's location.

Moreover, a recent study by Rubio-Fernández and Glucksberg (2012) further suggests that the ability of some human adults to take a third-person perspective on an agent's instrumental action can also be disrupted by being asked the where-prediction question by an experimenter. Monolingual and bilingual adults were asked the where-prediction question in a typical elicited-response false-belief task: 'Where will Sally look for her marble?' Rubio-Fernández and Glucksberg (2012) found, as expected, that all monolingual and bilingual adults who knew the location of the object correctly predicted that an agent with a false belief would look for it at the empty location. But using an eye-tracker, they further found that 100 ms after being asked the where-prediction question, 56.5% of the monolingual adults first looked at the actual (not the empty) location, while only 26.1% of the bilinguals did.[19] Thus, 2.5-year-olds fail the elicited-response false-belief Anne-Sally

[19] This finding showing that bilingual adults are better able than monolinguals to gaze at the empty location in anticipation of a mistaken agent's action fits with Kovacs and Mehler's (2009) finding that bilingual infants have stronger executive control resources than monolinguals. More recent findings based on eye-tracking reported by Rubio-Fernández (2013) further support the

task, but 25-month-olds correctly gaze in anticipation of the agent's action at the empty location where the agent falsely believes the object to be, in the absence of any verbal prompt (Southgate *et al.*, 2007). By contrast, all adults correctly predict where an agent with a false belief will look for an object when they are not under time pressure, but more than half of the monolingual adults first look at the actual location immediately after being asked the where-prediction question. This finding suggests that even adults may experience a reality bias exogenously generated by the where-prediction question before correctly answering the question.

4.2 The Referential Bias

While 2.5-year-olds seem perfectly able to track the contents of both the instrumental agent's false belief and the experimenter's communicative intention, from a detached third-person perspective (Scott *et al.*, 2012), their ability to correctly look at the empty location in anticipation of the mistaken agent's instrumental action (Southgate *et al.*, 2007) is disrupted if (but only if) they take the experimenter's where-prediction question to be addressed to them, not to be self-addressed by the experimenter to herself (He *et al.*, 2011). Even the ability of monolingual adults to gaze at the empty location in anticipation of the mistaken agent's action is impaired by the experimenter's where-prediction question. Conversely, when incited to act out the next action of a puppet (the Duplo girl) by an *open* question ('What happens next?')—instead of the where-prediction question—, a majority of 3-year-olds took the Duplo girl, who had a false belief about her food's location, to the empty location (in accordance with the content of the puppet's false belief), not to its actual location (Rubio-Fernández and Geurts, 2013).[20]

In their paper, Southgate *et al.* (2007) considered the possibility that 'the "where" question involved in most versions of this paradigm is prematurely interpreted by young children as referring to the location of the hidden object, rather than the location of the actor's subsequent actions'. Following their hint, we now turn to what we call the 'referential bias' generated by the pragmatics of the experimenter's asking the where-prediction question. As it will turn out, the referential bias rests on the different psychological processes required by complementary, but dissociable, aspects of elicited-response false-belief tasks.

From a detached third-person perspective, participants expect Sally (the mistaken agent) to look for her marble at the empty location. Their own knowledge of the

hypothesis that the where-prediction question can disrupt adults' ability to gaze at the empty location in anticipation of a mistaken agent's action.

[20] In addition to replacing the standard where-prediction question by an open question, the Duplo task also makes it easier for participants to keep track of the protagonist's perspective in three respects. First, since the experimenter moves the location of the bananas, participants can focus on the perspective of a single protagonist. Secondly, the protagonist never leaves the scene and is therefore visible all along. Thirdly, subtle communicative cues are designed to invite the participants to join the experimenter into the joint action of deceiving the protagonist.

object's location is consistent with their expectation that Sally will look for it at the empty location, as indicated by much evidence reviewed in this article. In particular, the evidence by He *et al.*, 2011 and by Scott *et al.*, 2012 shows that 2.5-year-olds are perfectly able to keep track of the mistaken agent's incorrect epistemic perspective on the object's location, even when the experimenter asks the where-prediction question, provided that the experimenter does not address the question to them, but either to herself or to some third-party. When they are asked an open question, the ability of 3-year-olds to keep track of the mistaken agent's epistemic perspective on the object's location does not seem disrupted either (Rubio-Fernández and Geurts, 2013). What is it about being the addressee of a where-prediction question by the experimenter that disrupts young children's ability to keep track of the mistaken agent's incorrect perspective on the object's location?

When the experimenter asks a where-prediction question, she is asking a WH-question, which she could not express unless her utterance contained a linguistic constituent referring to Sally's marble (i.e. 'Sally's marble'). In other words, a fundamental component of the experimenter's communicative action involves her referring to the object. In order to correctly answer the experimenter's where-prediction question, participants must form a thought involving a reference to the *location* of the mistaken agent's (i.e. Sally's) marble. While the constituent 'the location of Sally's marble' is not linguistically encoded in the experimenter's question, participants must in effect form a thought including a mental reference to the location of Sally's marble (where Sally will go to retrieve her marble). Now, the reference of 'the location of Sally's marble' can be fixed in participants' minds with respect to one of two competing epistemic perspectives: Sally's mistaken perspective (not shared by the children) or the epistemic perspective shared by participants and the experimenter with whose communicative action they are currently engaged from a second-person perspective.[21] While the latter determines the marble's actual location, the former determines the empty location. We surmise that the latter trumps the former.

In short, the referential bias itself rests on two components: first of all, participants must identify the object (the marble) in accordance with the experimenter's referential action. If and when they have done so, their attention is primed to the object's actual location. Secondly, the mistaken agent's incorrect epistemic perspective on the object's location is further downgraded (or overshadowed) in the participants' minds by the rival epistemic perspective on the object's location which they *share* with the experimenter who is addressing them.

This hypothesized two-tiered referential bias could be further tested in the following 2x2 experimental design: first, the experimenter could either refer to the

[21] As the evidence by Moll and Tomasello (2007) and Moll *et al.* (2007) shows, second-person cooperative engagement with a communicative agent significantly helps 14-month-olds to keep track of the communicative agent's epistemic familiarity with which toys they did and did not jointly attend to.

object by asking the where-prediction question ('Where will Sally look for her marble?') or fail to refer to the object by asking an open question ('What happens next?'). Secondly the experimenter could share either the participants' correct epistemic perspective on the object's location or the mistaken agent's incorrect epistemic perspective on the object's location. For example, the mistaken experimenter could leave the room with the mistaken agent and re-enter the room just before the mistaken agent does. The mistaken experimenter would thus share the mistaken agent's incorrect epistemic perspective on the object's location when she asks either the where-prediction question or the open question. In standard versions of the Sally-Anne task, the experimenter asks the where-prediction question (thereby referring to the object) and shares the participants' correct epistemic perspective. In the Duplo task (Rubio-Fernández and Geurts, 2013), the experimenter also shares the participants' correct epistemic perspective, but since she asks an open question, she fails to refer to the object. So the first factor has already been tested by Rubio-Fernández and Geurts (2013). In the two new cases, the experimenter would *share* the mistaken agent's *incorrect* epistemic perspective on the object's location (or be ignorant about it) and would either refer to the object (by asking the where-prediction question) or fail to refer to the object (by asking an open question). Testing these two new cases would enable us to disentangle to some extent the relative contribution of respectively the experimenter's reference to the object and her sharing the participants' correct epistemic perspective on the object's location in causing young children to point to the object's actual location. We expect that undermining the effects of either component of the referential bias would improve the performance of children of all relevant ages until they are at ceiling in standard elicited-response change-of-location false-belief tasks.[22]

4.3 The Cooperative Bias

The referential bias is pragmatically generated by the experimenter's communicative action and its primary effect is *epistemic*: it causes participants to inappropriately give priority to the correct epistemic perspective on the object's location which they share with the speaker, at the expense of the mistaken agent's incorrect perspective on the object's location. When, prior to the experimenter's question, participants are made aware that the success of an agent's instrumental action is being compromised by a false belief caused in her absence by a second agent, they may also spontaneously feel motivated to help the mistaken agent by a *cooperative bias*.

On the one hand, there is evidence that very young children do have a propensity to help mistaken agents achieve the goal of their instrumental action. The evidence gathered by Tomasello and colleagues (reviewed in Section 3) shows that very young

[22] Control questions in which they are asked about the toy's actual location prior to being asked the where-prediction question are also likely to disrupt participants' ability to keep track of the mistaken agent's epistemic perspective on the object's location.

human children have a spontaneous altruistic propensity to take a second-person perspective onto another's instrumental action and to help the agent achieve her goal. In particular, the findings by Buttelmann *et al.* (2009), Southgate *et al.* (2010) and Knudsen and Liszkowski (2012) are evidence that whether solicited or not, they correctly point to an object's actual location to help an agent achieve the goal of her instrumental action, whose success is compromised by her false belief about her desired object's location.

On the other hand and conversely, there is also evidence reported by Chandler *et al.* (1989) and Sullivan and Winner (1993), further assessed by Wellman *et al.*'s (2001) meta-analysis and to some extent confirmed by the findings based on the Duplo task (Rubio-Fernández and Geurts, 2013), that being actively involved in causing an instrumental agent to form a false belief helps 3-year-olds pass elicited-response change-of-location false-belief tasks. Being motivated to manipulate the instrumental agent's belief about the object's location seems to help young children attend to the resulting mistaken epistemic perspective on the object's location. In Rubio-Fernández and Geurts's (2013) experiment, when the experimenter checks the children's understanding that she is causing the Duplo girl to have a false belief, her action is described as 'secretive'. Arguably, the deceitful character of the experimenter's communicative action is sufficient to inhibit the propensity of 3-year-olds to help the mistaken agent when they are invited by the experimenter's open question to complete the Duplo girl's instrumental action.

While it is clear that participants cannot be motivated to deceive and to help one and the same individual at the same time, it is plausible to suppose that being actively involved in deception forces them to inhibit their altruistic tendency to help a mistaken agent by pointing to the object's actual location. If and when the inhibitory role of active deception is missing, participants may spontaneously be subject to the cooperative bias and endorse an advisory second-person perspective on the mistaken agent's instrumental action. The cooperative bias may thus contribute to turning the experimenter's *prediction* question ('Where *will* Sally look for her marble?') into the *normative* question ('Where *should* Sally look for her marble?'). Of course, the correct answer to the normative question is the object's actual location, not the empty location.

Another feature of the pragmatic context of the where-prediction question may also cause participants to turn the prediction question into a normative question. As the experimenter knows as much about both the object's actual location and the content of Sally's false belief about the object's location, participants might assume that the experimenter cannot be asking for information available to them and unavailable to the experimenter. Instead they might assume that the experimenter is testing their ability to tell where the mistaken agent *should* look for the object.

The cooperative bias explains Siegal and Beattie's (1991) and Surian and Leslie's (1999) findings because when the predictive meaning of the where-prediction question is linguistically enhanced by the insertion of the temporal adverb 'first', it is

pragmatically harder (if not downright impossible) to turn the question into a normative one. The cooperative bias hypothesis could be experimentally tested by a basic task manipulation that either counteracts young children's natural tendency to be cooperative or else allows children to pass a false-belief task while being helpful towards the mistaken agent. For example, consider a design in which the mistaken agent is shown to be a mean character (or a member of the participants' out-group) in contrast to a nice character (i.e. a member of the participants' in-group). We predict that error rates would decrease in the first condition relative to the second because children in the first condition would display a decreased propensity to cooperate with the mistaken agent. Furthermore, the cooperative bias may plausibly be construed as a resolution of the perspectival conflict created by the prediction question, whereby participants incorrectly take a second-person perspective towards both the experimenter and the mistaken agent. Participants might also resolve this perspectival conflict if offered the opportunity to covertly manipulate the mistaken agent's mind. For example, they might be prompted to surprise the mistaken agent by placing an unexpected pleasing object, which she will be happy to find at the location where she is likely to search for her own toy. We predict that this manipulation should significantly improve young children's performance in contrast to a neutral control condition.

Thus, the referential and the cooperative biases rest on different cognitive mechanisms. While the referential bias is jointly generated by the experimenter's referential communicative action and by the fact that the experimenter shares the child's correct epistemic perspective on the toy's actual location, its cognitive effect on the child is primarily epistemic. By contrast, while the cooperative bias reflects the child's spontaneous altruistic motivation to cooperate and help the mistaken agent achieve the goal of her instrumental action, in spite of her false belief, it causes the child to turn the prediction-question into a normative question. While further experimental investigation will help understand how they are related, the outcome of both biases is to dispose young children to point towards the toy's actual location.[23]

4.4 Answering Putative Challenges

So far, we have applied our pragmatic account exclusively to elicited-response *change-of-location* false belief tasks in which participants know the location of the toy which the mistaken agent is trying to retrieve. There are, however, other elicited-response false-belief tasks which we shall now briefly examine in light of our account. One important finding reported by Call and Tomasello (1999)

[23] As an anonymous reviewer interestingly remarks, our hypothesis that success at elicited-response false-belief tasks about an object's location requires overcoming the referential and the cooperative bias does not logically rest on acceptance of the processing-load account. As a result, advocates of cultural constructivism might also look forward to findings that test this pair of biases.

is a *non-verbal* false-belief task, in which participants who do *not* know the toy's location are required to determine it on the basis of the non-verbal testimony of a cooperative but mistaken agent. 4-year-olds have been shown to fail this task. Its significance lies in the fact that it is a non-verbal task so that its failure by young children may seem to challenge our pragmatic account.

In this task, children are requested to determine the location of a sticker that has been hidden in one of two opaque containers by one experimenter (hider *H*), with the help of a benevolent second experimenter (cooperative communicator *C*). *C* can, but the children cannot, see in which container *H* places the sticker. *C* leaves the room without informing the children and while *C* is away, *H* now switches the locations of the containers in front of the children. When *C* comes back, she non-verbally displays in front of the children the content of her *false* belief by marking one of the containers. The children must use the content of *C*'s (false) testimony to deduce the sticker's location.

Clearly, what is challenging for 4-year-olds in this task is *not* that they must take a third-person perspective on a mistaken agent's instrumental action, while taking a second-person perspective on the experimenter's communicative action. Instead, what is challenging is that it is a *non-verbal false communication* task (cf. Mascaro and Sperber, 2009). At t, children must first store the information that benevolent communicative agent *C* has a true belief without knowing its content. In other words, they know that *C* has a true belief, but they do not know which. Secondly, when at t + 1, they discover the content of *C*'s benevolent testimony, they must also *appraise* it as *false*. Thirdly, they must use their understanding of both the content of *C*'s testimony and the *logical* properties of a belief's *falsity* in order to deduce the location of the sticker. Finally, they must be able to understand that the testimony of a *benevolent* communicative agent could be *false*.

So-called unexpected-contents tasks, however, are different. In the so-called *Smarties* task, children are shown e.g. a familiar Smarties box whose external features suggest that it contains Smarties. When they are asked what is in the box ('What is in here?'), the majority of 3-year-olds answer: 'Smarties'. When the box is opened in front of them, the children discover that the box contains crayons, not Smarties. Finally they are asked what someone else not present when the box was opened will think is in the box upon seeing the box ('What will Mary think is in here, crayons or Smarties?'). The majority of 3-year-olds incorrectly answered that Mary will think that the box contains crayons (cf. Gopnik and Astington, 1988; Wellman *et al.*, 2001).

Recent evidence based on spontaneous-response tasks suggests that 2.5-year-olds are able to accurately represent the content of an agent's false belief about the unexpected content of a misleading box (cf. He *et al.*, 2011). In the agent's absence, an experimenter switched the content of a box of cheerios into a box of crayons and the content of a box of crayons into the box of cheerios. 2.5-year-olds looked reliably longer when the mistaken agent who loudly announced her desire to eat cheerios reached for the box containing cheerios than for the box containing crayons. If so, then they should also correctly *anticipate* a mistaken agent's action based on

the content of her false belief about the content of a perceptually misleading box.[24] This prediction could again be tested on the basis of in-group/out-group manipulations. For example, they should feel more inclined to *warn* an agent who wants to eat cheerios and falsely believes that the cheerios box contains cheerios if the agent belongs to the children's in-group than if she belongs to the out-group.

Clearly, the cooperative bias is irrelevant to explaining why 3-year-olds fail the elicited-response unexpected-contents tasks, but the experimenter's communicative action may nonetheless generate a *normative* bias in young children's minds for the following reason. In unexpected-contents tasks, young children are being *tricked*: they might take the experimenter's communicative action whereby the unexpected content of the Smarties box is displayed to them as a teaching demonstration and possibly as a warning signal about the deceptive character of perceptual appearances. If young children interpret the experimenter's teaching demonstration as a warning signal, then their primary motivation will be: not to be fooled again. If so, then they are likely to turn the experimenter's question about what they first thought was in the box into the *normative* question about what they *should* have first thought was in the box. Similarly, they are likely to turn the question about what another will think is in the box into the normative question about what another *should* think is in the box. Conversely, when young children are actively involved in causing another's false belief about unexpected content, they are of course less likely to turn the question about what another *will* think into the normative question about what another *should* think (Sullivan and Winner, 1993).

While young children may be motivated to change the questions about unexpected-contents into normative questions, the referential bias generated by the experimenter's reference to the box also applies to the pragmatics of the questions about unexpected-contents. When the experimenter asks the children what they first thought was in the box or what another will think is in the box, she refers to the box using the indexical 'here'. As a result, children, who share the experimenter's perspective on the current content of the box, must have opened an object-file for the box. Arguably, the box's current content (crayons) is a piece of information that is easily accessible in the children's object-file for the box. Furthermore, it is information shared by the children and the experimenter: while only the children, not the experimenter, earlier believed the box to contain Smarties, both the children and the experimenter now believe that the box contains crayons (not Smarties). In a nutshell, the experimenter's reference to the box activates the children's object-file for the box, which makes the content of her correct belief about the box's content (which the children share with the experimenter) salient to the children, at the expense of her own mistaken earlier epistemic perspective on

[24] Buttelmann *et al.* (2014) report that 18-month-olds succeed in a spontaneous-response unexpected-content false-belief task.

it, or at the expense of another's mistaken counterfactual epistemic perspective on it, which the experimenter does not share.[25]

5. Concluding Remarks

Much developmental psychology of the past thirty years or so has been devoted to false-belief tasks on the grounds that what could decisively show that an individual has mindreading abilities—a 'theory-of-mind', as Premack and Woodruff (1978) called it—is that she can track the content of another's psychological state different from her own. This research has uncovered discrepant experimental findings: on the basis of elicited-response tasks, it was widely believed that most children under 4 years of age are unable to track the contents of others' false beliefs. However, novel and startling evidence based on spontaneous-response tasks in the past ten years has also shown that before the end of their first year preverbal human infants expect others to act in accordance with the contents of their beliefs, including their false beliefs.

In this article, we opt for early-belief understanding as a resolution of the puzzle about the discrepant developmental findings because we think that major obstacles (both empirical and conceptual) stand on the way of cultural constructivism (cf. Section 1). Furthermore, there is also growing evidence (reviewed in Section 3.2) suggesting that by the age of 12-month-olds human infants are prepared both to fulfill the informative intention of an agent of an ostensive communicative action by acquiring some new belief and also to cause some new belief in others by their own ostensive pointing actions. If infants are prepared to acquire novel beliefs in response to another's ostensive communicative action and also to cause novel beliefs in others by their own ostensive communicative actions, then it seems as if they should be prepared to track the contents of others' epistemic states (including their false beliefs).

According to early-belief understanding, spontaneous-response false-belief tasks reveal the abilities of preverbal infants to track the contents of others' false beliefs and failure in elicited-response tasks need not reflect failure to track the

[25] For the purpose of examining and dismissing the potential role of executive functions in children's failure to respond to explicit prediction-questions ('Where will Sally look for her marble?') about the likely action of a mistaken agent, Perner *et al.* (2002) have asked children explanation-questions (e.g. 'Why did Sally look for her marble in the basket?'). They report that children failed equally to answer the explanation-questions and the prediction-questions. Since failure to answer explanation-questions is unlikely to reflect failure of executive functions, Perner *et al.* (2002) challenge the view that failure of executive functions might contribute to explain children's failure to correctly answer prediction-questions. Here, we can merely note that explanation-questions are likely to raise special pragmatic puzzles of their own for preschoolers. For one thing, preschoolers are used to ask, not to being asked, why-questions. Furthermore, answering a why-question requires the ability to eliminate irrelevant potential answers that are already known and taken for granted by the addressee.

contents of others' false beliefs. Moreover more and more evidence based on spontaneous-response false-belief tasks further shows that before the end of their second year, young children can track the contents of others' false beliefs about unexpected-contents, unobvious properties and object-identity.

The bulk of our article is devoted to discharging the burden of early-belief understanding by answering the question: what makes elicited-response false-belief tasks so challenging for most children younger than 4 years of age? We have argued that the processing-load account, which is meant to answer this question, is incomplete and we have offered a pragmatic account whose purpose is to fill the gaps of the processing-load account. In the article, we focus on elicited-response change-of-location false-belief tasks in which participants who know an object's location are asked to predict a mistaken agent's likely action. We also argue (at the end of the article) that our pragmatic framework sheds light on two other kinds of elicited-response false-belief tasks: unexpected-content false-belief tasks and change-of-location false-belief tasks, in which participants do not know an object's location and must determine it from the testimony of a mistaken agent.

Our pragmatic framework rests on the following twofold distinction: while humans can mind read the psychological states of agents of both instrumental actions and ostensive communicative actions, they can take either a third-person or a second-person perspective on either kind of action. To take a second-person perspective, but not a detached third-person perspective, on either a communicative or an instrumental action performed by another, is to recognize that one's own action (or response) is required for the success (or the failure) of the agent's action.

As we argue, the evidence based on spontaneous-response change-of-location false-belief tasks shows that when they take a detached third-person perspective on a mistaken agent's instrumental action, preverbal infants are able to track the contents of both her motivation and epistemic state (including her false belief). We have argued that in elicited-response change-of-location false-belief tasks, young children are confronted to two actions (not one): the mistaken agent's instrumental action and the experimenter's communicative action. Young children who know the toy's actual location are asked by the experimenter to predict where the mistaken agent will look for her toy. We have argued that most children under 4 years of age fail this task, not because they are unable to track the content of another's false belief, but instead because they must simultaneously take a third-person perspective with respect to the mistaken agent's instrumental action and a second-person perspective with respect to the experimenter's communicative action. They are overwhelmed by the tension between the two perspectives. In the context of elicited-response false-belief tasks about an object's location, this tension is likely to generate a referential and a cooperative bias, both of which are likely to highlight the object's actual location, at the expense of the empty location where the mistaken agent falsely believes the object to be. Future conceptual and experimental investigation of the early mindreading abilities whereby infants are caused to change their own minds by others' ostensive communicative actions and especially whereby they change the minds of others

by e.g. their own ostensive pointing is likely to shed more light on the distinction between second- and third-person perspectives.

K. Helming
Differential Psychology, Personality Psychology, and Diagnostics
University of Trier

B. Strickland and P. Jacob
Institut Jean Nicod
ENS/EHESS/CNRS

References

Apperly, I. 2011: *Mindreaders: The Cognitive Basis of 'Theory of Mind'*. New York: Psychology Press.

Apperly, I. and Butterfill, S. 2009: Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 116(4), 953–70.

Baillargeon, R., Scott, R. M. and He, Z. 2010: False-belief understanding in infants. *Trends in Cognitive Sciences*, 14, 110–8.

Baillargeon, R., Scott, R. M., He, Z., Sloane, S., Setoh, P., Jin, K., Wu, D. and Bian, L. 2015: Psychological and sociomoral reasoning in infancy. In M. Mikulincer, P.R. Shaver *et al.* (eds), *APA Handbook of Personality and Social Psychology: Vol.1. Attitudes and Social Cognition*. Washington, DC: American Psychological Association, 79–150.

Baron-Cohen, S., Leslie, A. and Frith, U. 1985: Does the autistic child have a 'theory of mind'? *Cognition*, 21, 37–46.

Bennett, J. 1978: Some remarks about concepts. *Behavioral and Brain Sciences*, 4, 557–60.

Birch, S. A. J. and Bloom, P. 2003: Children are cursed: an asymmetric bias in mental state attribution. *Psychological Science*, 14(3), 283–6.

Birch, S. A. J. and Bloom, P. 2004: Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences*, 8(6), 255–60.

Birch, S. A. J. and Bloom, P. 2007: The curse of knowledge in reasoning about false beliefs. *Psychological Science*, 18(5), 382–6.

Bloom, P. and German, T. 2000: Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77, B25–B31.

Buttelmann, D., Carpenter, M. and Tomasello, M. 2009: 18-month-olds infants show false-belief understanding in an active helping paradigm. *Cognition*, 112, 337–42.

Buttelmann, D., Over, H., Carpenter, M. and Tomasello, M. 2014: Eighteen-month-olds understand false beliefs in an unexpected-contents task. *Cognition*, 119, 120–6.

Butterfill, S. and Apperly, I. 2013a: How to construct a minimal theory of mind. *Mind & Language*, 28, 606–37.

Butterfill, S. and Apperly 2013b: Replies to three commentaries on minimal theory of mind. Online at http://philosophyofbrains.com/2013/11/11/symposium-on-butterfill-and-apperlys-how-to-construct-a-minimal-theory-of-mind-mind-language-28-5-606-63.aspx

Call, J. and Tomasello, M. 1999: A nonverbal false belief task: the performance of children and great apes. *Child Development*, 70, 381–95.

Call, J. and Tomasello, M. 2008: Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12(5), 187–92.

Carruthers, P. 2013: Mindreading in infancy. *Mind & Language*, 28, 141–72.

Chandler, M., Fritz, A. S. and Hala, S. 1989: Small scale deceit: deception as a maker of 2-, 3, and 4-year-olds early theory of mind. *Child Development*, 60, 1263–77.

Clements, W. A. and Perner, J. 1994: Implicit understanding of belief. *Cognitive Development*, 9, 377–95.

Csibra, G. 2010: Recognizing communicative intention in infancy. *Mind & Language*, 25, 141–68.

Csibra, G. and Gergely, G. 2009: Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148–53.

Csibra, G. and Volein, A. 2008: Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology*, 26, 1–11.

Darwall, S. 2006: *The Second-Person Standpoint*, Cambridge, MA: Harvard University Press.

Dennett, D.C. 1978: Beliefs about beliefs. *Behavioral and Brain Sciences*, 4, 568–70.

Devine, R.T. and Hughes, C. 2014: Relations between false belief understanding and executive function in early childhood: a meta-analysis. *Child Development*, 85(5), 1777–94.

Eilan, N. 2014: *The you turn. Philosophical Explorations*. Online at http://dx.doi.org/10.1080/13869795.2014.941910

Fodor, J.A. 1992: A theory of the child's theory of mind. *Cognition*, 44, 283–96.

Gergely, G. and Jacob, P. 2013: Reasoning about instrumental and communicative agency in human infancy. In J.B. Benson (Serial ed.) and F. Xu and T. Kushnir (vol. eds), *Rational Constructivism in Cognitive Development*. Waltham, MA: Elsevier Inc./ Academic Press, 59–94.

Gopnik, A. and Astington, J.W. 1988: Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59, 26–37.

Grice, H.P. 1969/1989: Utterer's meaning and intention. In his *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Hamlin, J. K., Wynn, K. and Bloom, P. 2007: Social evaluation by preverbal infants. *Nature*, 450, 557–9.

Harman, G. 1978: Studying the chimpanzee's theory of mind. *Behavioral and Brain Sciences*, 4, 576–77.

He, Z., Bolz, M. and Baillargeon, R. 2011: False-belief understanding in 2.5-year-olds: evidence from change-of-location and unexpected-contents violation-of-expectation tasks. *Developmental Science*, 14, 292–305.

He, Z., Bolz, M. and Baillargeon, R. 2012: 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology*, 30, 14–29.

Heyes, C. 1998: Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, 21, 101–148.

Heyes, C. 2014: False belief in infancy: a fresh look. *Developmental Science*, 17, 647–59.

Heyes, C. and Frith, C. 2014: The cultural evolution of mind reading. *Science*, 344, 1357–61.

Hutto, D. 2004: The limits of spectatorial psychology. *Mind & Language*, 19, 548–73.

Jacob, P. in press: A puzzle about belief-ascription. In B. Kaldis (ed.), *Mind and Society: Cognitive Science Meets the Philosophy of the Social Sciences*. Synthese Philosophy Library. Berlin: Springer.

Knudsen, B. and Liszkowski, U. 2012: Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Developmental Science*, 15, 113–22.

Kovacs, A.M. and Mehler, J. 2009: Cognitive gains in 7-month-old bilingual infants. *PNAS*, 106(16), 6556–60.

Kovacs, A. M., Erno, T. and Endress, A. D. 2010: The social sense: susceptibility to others' beliefs in human infants and adults. *Science*, 330, 1830–4.

Leslie, A. M. 1994: Pretending and believing: issues in the theory of ToMM. *Cognition*, 50, 211–38.

Leslie, A. M., Friedman, O. and German, T. P. 2004: Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, 8(12), 528–33.

Leslie, A. M., German, T. P. and Polizzi, P. 2005: Belief–desire reasoning as a process of selection. *Cognitive Psychology*, 50, 45–85.

Liszkowski, U., Carpenter, M., Striano, T. and Tomasello, M. 2006: 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, 7(2), 173–87.

Luo, Y. 2011: Do 10-month-old infants understand others' false beliefs? *Cognition*, 121, 289–98.

Mascaro, O. and Sperber, D. 2009: The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, 112, 367–80.

Matsui, T. and Miura, Y. 2008: Pro-social motive promotes early understanding of false belief. Available from *Nature Proceedings* <http://hdl.handle.net/10101/npre.2008.1695.1>

Meltzoff, A. N. and Brooks, R. 2008: Self-experience as a mechanism for learning about others: a training study in social cognition. *Developmental Psychology*, 44, 1257–65.

Milligan, K., Astington, J. W. and Dack, L. A. 2007: Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78, 622–46.

Moll, H. and Tomasello, M. 2007: How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, 43(2), 309–17.

Moll, H., Carpenter, M. and Tomasello, M. 2007: Fourteen-month-olds know what others experience only in joint engagement. *Developmental Science*, 10, 826–35.

Onishi, K. H. and Baillargeon, R. 2005: Do 15-month-olds understand false beliefs? *Science*, 308, 255–8.

Perner, J. Lang, B. and Kloo, D. 2002: Theory of mind and self-control: more than a common problem of inhibition. *Child Development*, 73(3), 752–67.

Perner, J. and Roessler, J. 2010: Teleology and causal reasoning in children's theory of mind. In J. Aguilar and A. Buckareff (eds), *Causing Human Action: New Perspectives on the Causal Theory of Action*. Cambridge, MA: MIT Press, 199–228.

Perner, J. and Roessler, J. 2012: From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, 16(10), 519–25.

Perner, J. and Ruffman, T. 2005: Infants' insight in to the mind: how deep? *Science*, 308, 214–6.

Piaget, J. 1954: *The Construction of Reality in the Child*. New York: Basic Books.

Premack, D. and Woodruff, G. 1978: Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4, 515–26.

Roessler, J. and Perner, J. 2013: Teleology: belief as perspective. In S. Baron-Cohen, S.H. Tager-Flusberg and M. Lombardo (eds), *Understanding Other Minds– Third Edition* (UOM-3). Oxford: Oxford University Press, 35–50.

Rubio-Fernández, P. 2013: Perspective tracking in progress: do not disturb. *Cognition*, 129, 264–73.

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

Rubio-Fernández, P. and Glucksberg, S. 2012: Reasoning about other people's beliefs: bilinguals have an advantage. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 38, 211–7.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T. and Vogeley, K. 2013: Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36, 393–462.

Scholl, B. J. and Leslie, A. M. 1999: Modularity, development, and 'theory of mind'. *Mind & Language*, 14, 131–53.

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

Scott, R. M. and Baillargeon, R. 2009: Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*, 80, 1172–96.

Scott, R. M., Baillargeon, R., Song, H. and Leslie, A. 2010: Attributing false beliefs about nonobvious properties at 18 months. *Cognitive Psychology*, 61, 366–95.

Scott, R. M., He, Z., Baillargeon, R. and Cummins, D. 2012: False-belief understanding in 2.5-year-olds: evidence from two novel verbal spontaneous-response tasks. *Developmental Science*, 15, 181–93.

Senju, A. and Csibra, G. 2008: Gaze following in human infants depends on communicative signals. *Current Biology*, 18, 668–71.

Senju, A., Southgate, V., Snape, C., Leonard, M. and Csibra, G. 2011: Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science*, 22, 878–80.

Siegal, M. and Beattie, K. 1991: Where to look first for children's knowledge of false beliefs. *Cognition*, 38, 1–12.

Southgate, V., Senju, A. and Csibra, G. 2007: Action anticipation through attribution of false belief by two-year-olds. *Psychological Science*, 18, 587–92.

Southgate, V., Chevallier, C. and Csibra, G. 2010: Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, 13, 907–12.

Southgate, V. and Vernetti, A. 2014: Belief-based action prediction in preverbal infants. *Cognition*, 130, 1–10.

Spaulding, S. 2013: Commentary on 'How to construct a minimal theory of mind'. Online at http://philosophyofbrains.com/2013/11/11/symposium-on-butterfill-and-apperlys-how-to-construct-a-minimal-theory-of-mind-mind-language-28-5-606-63.aspx

Sperber, D. and Wilson, D. 1986: *Relevance, Communication and Cognition*. Cambridge, MA: Harvard University Press.

Strickland, B. and Jacob, P. 2015: Why reading minds is not like reading words. Online at http://www.cognitionandculture.net/home/blog/44-pierre-jacobs-blog/2669-why-reading-minds-is-not-like-reading-words

Sullivan, K. and Winner, E. 1993: Three-year-olds' understanding of mental states: the influence of trickery. *Journal of Experimental Child Psychology*, 56, 135–48.

Surian, L., Caldi, S. and Sperber, D. 2007: Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18, 580–6.

Surian, L. and Leslie, A. 1999: Competence and performance in false belief understanding: a comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology*, 17, 141–55.

Tomasello, M. 2008: *Origins of Human Communication*. Cambridge, MA: MIT Press.

Tomasello, M. 2009: *Why We Cooperate.* A Boston Review Book. Cambridge MA: MIT Press.

Tomasello, M. 2014: *A Natural History of Human Thinking*. Cambridge, MA: Harvard University Press.

Topal, J., Gergely, G., Miklosi, A., Erdohegyi, A. and Csibra, G. 2008: Infant perseverative errors are induced by pragmatic misinterpretation. *Science,* 321, 1831–4.

Warnecken, F. and Tomasello, M. 2006: Altruistic helping in human infants and young chimpanzees. *Science*, 311, 1301–3.

Warneken, F. and Tomasello, M. 2007: Helping and cooperation at 14 months of age. *Infancy*, 11, 3, 271–94.

Warneken, F. and Tomasello, M. 2009: Varieties of altruism in children and chimpanzees. *Trends in Cognitive Sciences*, 13(9), 307–402.

Wellman, H.M., Cross, D. and Watson, J. 2001: Meta-analysis of theory of mind development: the truth about false belief. *Child Development*, 72, 655–84.

Wilson, D. and Sperber, D. 2004: Relevance theory. In L. Horn and G. Ward (eds), *The Handbook of Pragmatics*. Oxford: Blackwell.

Wimmer, H. and Perner, J. 1983: Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–28.

Zawidzki, T. 2013: Commentary on Butterfill and Apperly's 'How to construct a minimal theory of mind'. Online at http://philosophyofbrains.com/2013/11/11/symposium-on-butterfill-and-apperlys-how-to-construct-a-minimal-theory-of-mind-mind-language-28-5-606-63.aspx