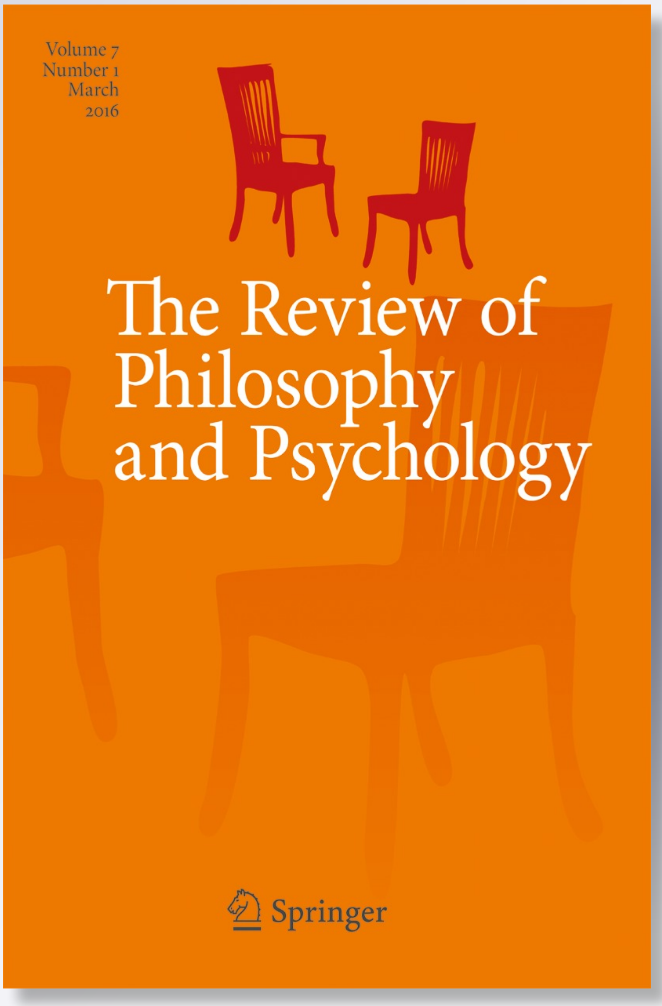
Two Systems for Mindreading?

Peter Carruthers

Review of Philosophy and Psychology

ISSN 1878-5158
Volume 7
Number 1

Rev.Phil.Psych. (2016) 7:141-162
DOI 10.1007/s13164-015-0259-y



Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com”.

Author's personal copy

Rev.Phil.Psych. (2016) 7:141–162
DOI 10.1007/s13164-015-0259-y

Two Systems for Mindreading?

Peter Carruthers[1]

Published online: 15 April 2015
© Springer Science+Business Media Dordrecht 2015

Abstract A number of two-systems accounts have been proposed to explain the apparent discrepancy between infants' early success in nonverbal mindreading tasks, on the one hand, and the failures of children younger than four to pass verbally-mediated false-belief tasks, on the other. Many of these accounts have not been empirically fruitful. This paper focuses, in contrast, on the two-systems proposal put forward by Ian Apperly and colleagues (Apperly & Butterfill, *Psychological Review*, 116, 953–970 2009; Apperly, 2011; Butterfill & Apperly, *Mind & Language*, 28, 606–637 2013). This has issued in a number of new findings (Apperly et al., *Psychological Science*, 17, 841–844 2006a; Back & Apperly, *Cognition*, 115, 54–70 2010; Qureshi et al., *Cognition*, 117, 230–236 2010; Samson et al., *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1255–1266 2010; Schneider et al., *Journal of Experimental Psychology: General*, 141, 433–438 2012a, *Psychological Science*, 23, 842–847 2012b, *Journal of Experimental Psychology: General*, 141, 433–438 2014a, *Psychological Science*, 23, 842–847 2014b; Surtees & Apperly, *Child Development*. 83, 452–460 2012; Surtees et al., *British Journal of Developmental Psychology* 30, 75–86 2012, *Cognition*, 129, 426–438 2013; Low & Watts, *Psychological Science*, 24, 305–311 2013; Low et al., *Child Development*, 85, 1519–1534 2014). The present paper shows that the theoretical arguments offered in support of Apperly's account are nevertheless unconvincing, and that the data can be explained in other terms. A better view is that there is just a single mindreading system that exists throughout, but which undergoes gradual conceptual enrichment through infancy and childhood. This system can be used in ways that do, or do not, draw on executive resources (including targeted searches of long-term memory) and/or working memory (such as visually rotating an image to figure out what someone else sees).

✉ Peter Carruthers
pcarruth@umd.edu

[1] University of Maryland, College Park, MD 20742, USA

Author's personal copy

1 Explaining the Gap: Two Theories

Since 2005 a wealth of data has been produced by multiple laboratories using a variety of different nonverbal methods, controls, and dependent measures indicating that infants between the ages of 6 and 18 months are capable of representing and reasoning about the false beliefs of other agents (Onishi and Baillargeon 2005; Southgate et al. 2007; Surian et al. 2007; Song et al. 2008; Buttelmann et al. 2009; Poulin-Dubois and Chow 2009; Scott and Baillargeon 2009; Kovács et al. 2010; Scott et al. 2010; Southgate et al. 2010; Träuble et al. 2010; Luo 2011; Senju et al. 2011; Knudsen and Liszkowski 2012; Yott and Poulin-Dubois 2012; Baillargeon et al. 2013; Barrett et al. 2013; Buttelmann et al. 2014; Southgate and Vernetti 2014; Buttelmann et al. 2015; Scott et al. 2015). In contrast, the results of hundreds of verbally-based tests with older children suggest that children under 4 years of age are incapable of such reasoning (Wellman et al. 2001). Various attempts have been made to explain the discrepancy. Some have argued that the infancy data can be accounted for without supposing that infants are capable of representing mental states at all, suggesting that their behavior can be explained in terms of low-level sensory associations (Perner and Ruffman 2005; Heyes 2014) or behavior-rules of some sort (Perner 2010; Gallagher and Povinelli 2012). But these approaches have not proven empirically fruitful: no new discoveries have resulted. And in every case where a determinate proposal of this sort has been tested and controlled for, it has been refuted. In any case, I will not discuss such accounts here. This paper will focus on the contrast between two-systems views of the sort proposed by Apperly and colleagues, and one-system accounts that attribute three-year-olds' failures in verbal false-belief tasks to problems of performance, not lack of competence.

1.1 The Two-Systems Account

While others have put forward two-systems views of various kinds (De Bruin and Newen 2012), I will focus here on the account provided by Apperly and colleagues, because of its clarity and its impact on the field (Apperly and Butterfill 2009; Apperly 2011; Butterfill and Apperly 2013). This claims that humans have *two* systems for representing and reasoning about the mental states of other agents, one of which is early-developing and is available in infancy, and the other of which is slower to develop and depends on language.[1] The former is said to explain the infancy data, whereas it is the absence of the latter that explains children's failures in verbal false-belief tasks before the age of four. Both are said to exist alongside one another in adults. The early-developing system is said to be fast and automatic in its operations, whereas the later-developing one is said to be slower and more effortful, but also more flexible. There is now a significant body of data collected with adults that is said to support the existence of these two systems (Apperly et al. 2006a; Back and Apperly 2010; Qureshi et al. 2010; Samson et al. 2010; Schneider et al. 2012a, 2012b, 2014a, 2014b; Surtees and Apperly 2012; Surtees et al. 2012, 2013; Low and Watts 2013; Low et al. 2014). This will be discussed in Section 3.

[1] The dependence here is intended to be developmental, not constitutive, since severely aphasic people can have normal mindreading abilities (Apperly et al. 2006b).

Author's personal copy

It is said that the early-developing system is fast because it is largely encapsulated from the rest of cognition (and is thus comparatively inflexible). Moreover, this system is said to employ a distinct set of concepts from those used by the later-developing one. Specifically, it has concepts that can *track* goals, perceptions, and beliefs, but without representing them as such. Rather, it conceptualizes goals in terms of the teleological *functions* of action (that is, goals are states of the world at which actions are aimed), and it conceptualizes beliefs as the results of *registerings* of past facts (where these are individuated extensionally, composed of worldly individuals, properties, and relations). It cannot attribute propositional attitudes (nor quantified thoughts) to agents, and it is insensitive to the aspectual nature of belief and perception (Butterfill and Apperly 2013). In contrast, the later-developing system represents goals and beliefs as such, can handle propositional attitudes of all sorts, and is sensitive to the aspectual character of mental representation generally.

This account has the resources to explain how infants can consistently display evidence of mental-state understanding across a range of nonverbal tasks. This is because they are sensitive to people's goals and registerings, and these can be used to issue in simple forms of prediction and explanation of action. The account can also explain *some* of the data regarding three-year-olds' failures in verbally mediated tasks, specifically cases where children are asked to report on their own or other people's beliefs. For by hypothesis the nearest available concept (a *registering*) in the early-developing system (onto which the linguistic term "believe" might otherwise be fast-mapped, hence ensuring success) is encapsulated from language and other systems of the mind. So when learning terms like "believe" and "think" children are forced to construct a concept *de novo*, using general-purpose learning mechanisms, and relying on linguistic and other cues to do so. This concept is not yet fully available to three-year-olds, it can be claimed.

However, the account cannot so easily explain young children's failures in tasks where they are asked, not what someone thinks or believes, but rather what they will say or do, or where they will look. This is because it appears from the infancy data that the *output* of the early-developing system is not encapsulated from executive decision, since infants can take executively-controlled actions such as helping to open a box or reaching for a requested object on the basis of their beliefs about other people's false registerings (Buttelmann et al. 2009, 2014; Southgate et al. 2010). If three-year-olds can figure out what someone is trying to achieve given that they have falsely registered something, and can take a decision to help accordingly (thus relying on the output of the early developing system in their decision making), then why is it that they cannot rely on that output when saying what someone who has falsely registered something, and has a known goal, will do? It cannot be that the early-developing system can output goal-attributions but not action-predictions, since we have known for a while that three-year-olds can act appropriately for predicted actions that are based on false registerings, for instance by placing a pillow at the bottom of the slide that an agent with a false belief is predicted to come down (Clements and Perner 1994).[2]

Perhaps it can be claimed that the output of the early-developing system, while available to executive processes generally, is for some reason encapsulated from the language system. One problem with this suggestion is that it requires us to draw a sharp

[2] Note that these same three-year-olds fail *verbal* tests of their false-belief understanding.

Author's personal copy

distinction between the executive processes involved in a decision to act, quite generally, and those implicated in decisions to say one thing rather than another. This is at best ad hoc, and seems implausible.[3] Moreover, we know that in adults the output of the early-developing system is available to language production processes, since its putative outputs can at least impact reaction-times and error-patterns in people's reports of what they or another person can see (Qureshi et al. 2010; Samson et al. 2010). Why, then, are those outputs not so available in young children? Of course it can be claimed that the outputs of the early-developing system only *become* available to influence language production later in childhood, but this, too, seems ad hoc, and is not independently plausible.

The explanation that the two-systems account gives of the discrepancy between infants' success in nonverbal tasks and three-year-olds' failures in verbally-mediated tasks is not fully satisfying, then. Correspondingly more weight will therefore need to be placed on the other arguments offered in support of such an account. These will be considered in Sections 2 and 3.

1.2 The One-System Account

A number of different one-system accounts have been proposed (Leslie et al. 2004; Baillargeon et al. 2010; Carruthers 2013).[4] Here I will sketch what I take to be the most plausible version of such a view. On this account, infants start with a set of conceptual primitives (perhaps THINKS, LIKES, IS AWARE OF, and maybe TELLS), together with some simple rules for determining the application of these concepts. For example, if an infant sees a toy placed into a box in the presence of another agent, the infant encodes THE AGENT IS *AWARE* OF THE TOY GOING INTO THE BOX, and computes and thereafter remembers THE AGENT *THINKS* THE TOY IS IN THE BOX.[5]

At this early stage infants' concepts of true and false belief are implicit in their procedures for updating thought-attributions. Thus if the target object is placed in the box while the agent is present, the infant stores THE AGENT THINKS THE TOY IS IN THE BOX, and when the object is subsequently moved to a cupboard while the agent is absent, this representation is left unchanged (thus rendering it false); whereas if the agent is present when the item is

[3] Note, too, that no such distinction applies in the case of the other early-developing system that Apperly and colleagues take as their model, namely the approximate number system. For the output of this system is available not only to inform children's decisions about which of two bowls of treats they should choose, for example, but also enables them to *say* which bowl contains more.

[4] Baillargeon et al. (2010) actually describe their view as a *two*-systems account. They think that there are two early-developing mindreading systems, one of which develops slightly in advance of the other. The first handles attributions of goals, perceptual access, and knowledge or ignorance. The second deals with belief and false-belief, together with misleading appearances. But this is better characterized as a one-system view where the mindreading system is made up of distinct *components*. For there is no suggestion of radical conceptual change with further development in childhood, and the two components are not appealed to when explaining why three-year-olds fail verbal tasks. Nor is there any suggestion of distinct systems with very different processing profiles existing alongside one another in adulthood. I am myself quite happy to allow that the mindreading system might divide into a number of dissociable components. Note, however, that the evidence of false-belief understanding at 6 or 7 months of age (Kovács et al. 2010; Southgate and Vernetti 2014) suggests at least that all of the various components come online at about the same early age.

[5] Throughout I use small capitals to signify mental representations (as opposed to the contents of those representations). I assume that these representations are abstract and amodal in nature, and can be structured out of component concepts. But it is not presupposed that the resulting thoughts are language-like in nature (they might be map-like or diagram-like, for example). See Carruthers (2015b) for defense of these assumptions.

Author's personal copy

moved, the representation is updated to THE AGENT THINKS THE TOY IS IN THE CUPBOARD. Later, explicit concepts of truth and falsity can be introduced through a pair of straightforward definitions: An agent's belief that P is true = the agent thinks that P, & P; an agent's belief that P is false = the agent thinks that P, & not-P. The infant also continues to learn about the factors that can influence the application of mental-state concepts, such as wearing blindfolds, events happening in dim light, etc., and perhaps differentiates between different modes of sensory access (if these differences were not already among the conceptual primitives).

Importantly, on this account, the transition from infancy to later childhood is one of continuous conceptual enrichment and change. Concepts become differentiated and elaborated, and new principles for ascribing mental states to people are learned, as are domain-specific executive procedures, such as mentally rotating a visual image when figuring out what can be seen from someone else's perspective. But no radically new concepts are introduced (in the sense of Carey 2009), and there are never two separate systems for mindreading existing alongside one another with distinct conceptual resources. On the contrary, the early-developing system continues to develop, adding and differentiating concepts, adding principles of attribution and inference, thus gradually transforming into the mindreading system employed by adults. We can suppose that the system also becomes more efficient in its operation over time, and interacts more robustly with surrounding mental faculties.[6]

How is the one-system account to explain the failures of three-year-olds in verbal false-belief tasks? If infants are already capable of representing and reasoning about false beliefs, why is it that three-year-olds systematically answer wrongly when asked about false beliefs, or when asked to articulate predictions based on attributed false beliefs? While a number of different one-system explanations have been offered (e.g., Baillargeon et al. 2010), my own suggestion is that verbal false-belief tasks are effectively triple-mindreading tasks. The mindreading system will be engaged when interpreting the target agents and events, of course, attributing goals and beliefs, and forming expectations for behavior. But in addition it will be involving in processing the speech of the experimenter, helping to figure out the speaker's communicative intentions, especially at the point when the speaker shifts perspective from that of story-teller to asking the child direct questions. And then thirdly the mindreading system will need to assist in the production of speech or other communicative actions (like pointing), helping to evaluate which of various potential actions would have the desired effect on the mind of the questioner. Engaging in all three things at once, or over a short period of time, will require the child to switch back and forth between different perspectives (now the protagonist, now the experimenter, now one's own), bearing in mind the output of one round of mindreading computation while undertaking another, and then accessing the former at the appropriate time.[7]

[6] Consistent with these suggestions, Saxe et al. (2009) find increasing response-selectivity with age in the right temporo-parietal junction, which is known to be a crucial component of the mindreading system in adults.

[7] Is it a problem for this account that there are wholly non-verbal false-belief tasks that even 4-year-olds will fail? I suggest not. Indeed, when one looks at the details of such tasks, the incremental-development account is, if anything, further supported. Thus in Call and Tomasello (1999) the children would have had to represent that the agent had *some* false belief without knowing what that belief was. This can't be handled just by computing what someone believes and then failing to update when things change in their absence. On the contrary, it seems to require an explicit concept of false belief, as well as a capacity to quantify over false beliefs.

Author's personal copy

It is small wonder, then, that success in passing verbal false-belief tasks should correlate with executive function (Carlson et al. 2002, 2004; Kloo and Perner 2003) and linguistic ability (Astington and Jenkins 1999; Milligan et al. 2007), nor that it should improve with additional communicative experience (Perner et al. 1994; Kovács 2009) and be held back by the absence of such experience (Peterson and Siegal 1995; Woolfe et al. 2002). For success will require a subtle interplay between executive-function / working-memory systems, the language faculty, and the mindreading system itself. One might expect that the efficiency of each of these systems would improve with development and be strengthened by social and communicative experience, as would the connections between them.

Can this one-system account explain three-year-olds' *systematic* failures in verbal false-belief tasks, however? For in general young children do not answer randomly (suggesting confusion) but offer a reality-based response. In fact this is what we should expect if any one of the three components collapses under load. If a representation of the false belief of the agent is lost, then the language production process is likely to default to the next-most-relevant answer, which is what the children themselves would believe or do. If the experimenter's question is incorrectly interpreted, the next-most-likely interpretation will be a question about what is really the case, or what should really be done. And if the production process goes awry, again the most likely error will be to select the child's own belief or own likely action as the target for report.

This explanation of the gap between infant success in spontaneous tasks and children's later success in verbal tasks is consistent with, but somewhat broader than, the suggestion made by Baillargeon et al. (2010). They say that it is the late maturation of the neural pathways connecting the frontal lobes with the temporo-parietal junction that is responsible for young children's failure in language-involving false-belief tasks. For although all active-response tasks will implicate these connections to some degree, one should expect that language-based tasks would place much greater demands on the connections between mindreading and executive function, for the reasons sketched above. I differ, however, in suggesting that it is a three-way maturation that is required, involving not only mindreading and executive function, but also the language system.

The account sketched here is also consistent with the analysis and supporting evidence provided by Rubio-Fernández and Geurts (2013). They reason that standard verbal false-belief tasks disrupt the child's perspective taking. For example, there is generally more than one agent involved in the target story (Sally and Ann, say), and at a certain point the experimenter will shift from narrative mode to direct questioning of the child. Moreover, the questions themselves generally call attention to the true facts of the situation (as known by the child), making those representations more salient, and hence more likely to win the competition for expression in speech. Rubio-Fernández & Geurts devised a simple (but nevertheless largely verbal) task that minimized these perspective shifts, while also encouraging the child to use the Duplo character to act out the denouement of the story (rather than answering a direct question). They found that three-year-olds reliably passed.

A prediction of the present account is that one should be able to vary the difficulty of false-belief tasks systematically by ratcheting up, or down, the amount of language involved (either in comprehension, or production, or both), and/or by varying the demands placed on executive function (for example, if crucial facts need to be accessed and recalled to generate a prediction after attention has been drawn elsewhere), and/or

Author's personal copy

by varying the number of perspective changes (or more generally the number and complexity of the mental states) that the child needs to keep track of. While this set of predictions has not yet been systematically tested, it is consistent with results already in the literature. For example, Scott et al. (2012) show that 2.5-year-olds can pass false-belief tasks in which the materials are presented verbally, but where the dependent measure is not a response to a direct question but rather spontaneous looking behavior. and Buttelmann et al. (2009) find that 16-month-olds fail the same tasks that 18-month-olds pass, although 16-month-old infants have passed other kinds of false-belief task. The difference may be that the experiment required some processing of experimenter speech in addition to tracking and reasoning about the beliefs and goals of the target agent.

1.3 Simplicity Considerations

One motive for resisting the idea that infants can attribute propositional attitudes to others is a general skepticism that infants are capable of the cognitive demands required for such thinking. Many feel that attributing to infants a simpler set of capacities is preferable. They think that we should attribute the simplest, least cognitively sophisticated, capacities to infants that are consistent with the data. But this can only be on the assumption that the resulting theory is explanatorily adequate, of course. And as we have seen, it is doubtful whether the two-systems account can fully explain why three-year-olds should generally fail all forms of verbal false-belief task.

Moreover, the simplicity of the capacities we attribute to infants has a bearing on the simplicity of our overall theories of social development. A simpler theory at the initial stage may require us to postulate two distinct mindreading systems at the endpoint of development (as do Apperly and colleagues). In addition, a simpler starting-theory might require us to postulate some kind of acquisition process that can issue in radical conceptual change through the early years (introducing new conceptual primitives as opposed to conservative extensions of existing knowledge). It is the simplicity of our overall theory of a domain that counts, not the simplicity of theories designed to explain one particular body of evidence (the infancy data) within the domain.

Indeed, I suggest that it is the extra complexity entailed by postulating two distinct mindreading systems that needs to be justified. For as we have seen, this does not appear to be necessary (and nor is it sufficient) to explain the three-year gap between infants' competence with false-belief and later success in verbal false-belief tasks. As a result, the two-systems account will need to seek support from other considerations, including the evidence of two separable mindreading systems obtained with adults.

2 Representing Beliefs as Such

Butterfill and Apperly (2013) employ a number of theoretical arguments for the conclusion that infants aren't ascribing beliefs *as such* (and hence for the conclusion that the early-developing system contains a distinct set of concepts). One is that infants as yet have no awareness (they say) that distinct belief-contents can be used to represent the same state of affairs. Another is that infants have no awareness of the *normative* character of belief in relation to the available evidence. And a third is that it is

Author's personal copy

implausible that infants could represent all of the complexity of belief, given that beliefs have contents that are indefinitely complex and multiply nestable, while having numerous functional connections with mental states of other sorts. It is also held to be implausible that adults could deploy concepts of this complexity in a swift and automatic fashion—leading, again, to the need for a distinct set of concepts. I will take these points in turn.

2.1 Representing *Aspect*: the Evidence

Butterfill and Apperly (2013) claim that the early-developing system can represent that people register some facts but not others, but maintain that it is incapable of representing that people can be aware of some *aspects of* or *ways of representing* a fact but not others. Yet they claim it is essential for possession of the concept *belief* that one understand such possibilities. Roughly speaking, one can think of these as *Oedipus cases*. Oedipus can believe that Jocasta is only 5 years his senior while believing (obviously) that his mother is more than 5 years his senior, even though Jocasta is his mother. He holds one belief when the woman in question is represented under the aspect *Jocasta* and a contradictory belief when she is represented under the aspect *my mother*. Butterfill & Apperly claim that infants are incapable of understanding such cases, and hence lack the concept *belief*.

In making this claim Butterfill & Apperly are forced to reject the implications of the data provided by Scott and Baillargeon (2009), which seem to suggest that 18-month-old infants can ascribe false beliefs about object identity. The infants were exposed to an agent interacting with two identical-looking penguin-shaped objects, one of which could be separated into two parts and the other of which could not. The divisible penguin was always presented to the agent in its divided state, and the agent always used the divisible penguin as a hiding place for a key. In the test condition the agent saw the divisible penguin in its *conjoined* state, with the other penguin hidden underneath an opaque cover. The infants were surprised when the agent reached for the divisible penguin, and were not surprised when the agent reached for the penguin under the cover, seemingly because they ascribed to the agent the false belief that the uncovered penguin was the inseparable penguin, as well as the false belief that the divisible penguin was hidden under the cover. If so, then infants understand that the agent believes one thing of the penguin under one aspect (that it is penguin-shaped, say) while believing something else under another (that it is indivisible).

Butterfill and Apperly (2013) argue that this evidence fails to show that the infants ascribe false beliefs about object *identity*. Rather, it is consistent with the infants ascribing false beliefs about *types* of object. In particular, they might ascribe to the agent the belief that the uncovered penguin is *an* indivisible penguin, combined with the belief that *a* divisible penguin is always present, therefore expecting the agent to reach underneath the cover to retrieve *a* divisible penguin in which to hide the key. This is, indeed, a possible interpretation of the results. But notice that it nevertheless involves attributing to infants a capacity to ascribe quantified propositions to agents. (For example, "She thinks that a divisible penguin is always present.") Yet this, too, is supposed to be beyond the reach of the early-developing system. So either way, it appears that the early-developing system has greater sophistication than Butterfill & Apperly are prepared to allow.

Author's personal copy

Moreover, Butterfill and Apperly (2013) take note of the results of Song and Baillargeon (2008), which show that 15-month-old infants can predict actions on the basis of how things *appear* to agents who are ignorant of their true nature. (Specifically, infants predict that an agent searching for a blue-haired doll will open a box that has tufts of blue hair sticking out from beneath the lid, even though the infant knows that the box contains another object instead.) Butterfill & Apperly also take note of the results of Luo and Beck (2010), which show that infants expect an agent who is known to prefer red things will reach for a screen that has a red side facing the agent and a green one facing the infant, but not to reach for a screen with a green side facing the agent but a red one facing the infant.[8]

Butterfill & Apperly concede, in light of these results, that when infants generate expectations about people's behavior they are sensitive to the way things *look* to those agents, and further evidence for this conclusion continues to accumulate. Thus Buttelmann et al. (2015) use minimally-verbal versions of a set of classic appearance–reality tasks with active-helping as the output measure, showing that 18-month-old infants reliably pass. Moreover, Scott et al. (2015) show that 17-month-old infants expect an agent to be deceived by a thief who has substituted an identical-looking object for the goal object, thus seeming to attribute to the agent the belief that that the substitute object *is* the goal object.

Granting that infants understand how things look to other agents comes very close, at least, to conceding that infants understand the aspectuality of the mind. What is certainly true is that developmental psychologists have traditionally thought that a capacity to understand misleading appearances requires the same theoretical and conceptual resources as a capacity to understand both the representational nature of the mind and false belief as such (Gopnik and Meltzoff 1997). The burden of proof surely falls on defenders of the two-systems account to show how the conceptual demands of these misleading-appearance tasks are really quite different from those required for understanding the aspectual nature of belief.

What of supportive evidence in the other direction, however? There are just two published experiments using nonverbal measures suggesting that neither three-year-old children nor adults will automatically compute, nor derive expectations from, the false identity-beliefs of a target agent (Low and Watts 2013; Low et al. 2014). Yet in the same conditions both four-year-olds and adults were able to give correct verbal answers to questions. These data are said to show both the existence and signature limits of the early-developing system, and its continued existence alongside the later-developing verbally-mediated one.

Low and Watts (2013) first familiarized participants with the fact that an agent reliably wanted objects of one color rather than another (red rather than blue, say). Participants were also familiarized with the fact that when the doors flashed, the agent would reach though one of two available doors nearest to where she expected the object of the desired color to be. They then saw the agent watching while an object that appeared to the participants to be blue moved from one opaque box to another, where it swiveled back and forth to reveal to the participant (but not the target agent) that it was blue on one side but red the other. So the agent would have seen what she took to be a *red* object moving to that box. The same object then moved back to the other opaque

[8] Moll and Meltzoff (2011) provide a similar demonstration with three-year-old children.

Author's personal copy

container, but this time with its blue side facing the *agent*, who should therefore expect that the desired red object would still be in the first container, and should consequently reach for that position when the doors flash. In this experiment (in contrast with a control change-of-location false-belief task) neither three-year-olds, nor four-year-olds, nor adults looked in anticipation at the correct location.

Notice, however, that the inferential demands made of participants in this experiment are much greater than they are in a standard change-of-location task. In the latter, participants just need to encode a belief of the form THE AGENT THINKS THE OBJECT IS IN LOCATION X, and then forebear from updating this belief in the false-belief condition when the object is moved to location *y* in the agent's absence. In the identity version of the task, in contrast, participants will first encode (before they learn that the object is double-sided) THE AGENT THINKS THERE IS A BLUE OBJECT IN THE RIGHT-HAND BOX. When they see that the object is double-sided, they then need to erase this belief, having inferred that the agent would have seen what she would take to be a *red* object moving to that location, replacing it with THE AGENT THINKS THERE IS A RED OBJECT IN THE RIGHT-HAND BOX. Note that this inference would probably require executively controlled use of working memory, involving visual rotation of a memory image of the moving object in the original sequence in order to imagine what the agent would have seen. Then when the object moves back to the other box with its red side facing the participant, the participant would need to use visual rotation once again to infer that the agent would see what she would believe to be a *blue* object leaving the right-hand location (and would therefore expect the desired red object still to be there). All this is plainly much more cognitively demanding than a standard change-of-location task. It is little wonder that people should fail to engage the executively-controlled procedures necessary for success in the absence of any overt task or question.

Similar points apply to the findings of Low et al. (2014). Here children and adults watched while a figure rotated to reveal (to them, but not to the target agent, who was not present at this point) that what looked like a duck on one side had the appearance of a rabbit on the other. Although they had earlier been familiarized with the agent's goal of reaching through one of two doors to feed the animal the appropriate sort of food, they failed to anticipate in their looking that the agent would reach for the carrots, say (seeing what they would take to be a rabbit) when the figure appeared to the observer to be a duck. Here, too, if observers are to generate a prediction they would either need to rotate their image of the figure to compute what the agent would see, or they would need to access a memory of what the other side of the figure looked like. Either way, significant executive resources would surely be required.

Taken all together, then, I suggest that the data from infants considered thus far is more supportive of a one-system account than it is of the two-systems view. There is evidence that infants *can* represent the aspectual nature of belief (or something quite close to it), and apparent evidence to the contrary can be explained in terms of the executive demands of the tasks.

2.2 Representing *Aspect*: the Implications

Let us suppose, however, that Butterfill and Apperly (2013) are correct that infants are incapable of representing the aspectuality of belief; and suppose we grant, too, that such capacities are essential for possessing the concept of *belief*. Would this support a two-

Author's personal copy

systems account? It would not. Asking in this way whether or not infants possess the concept of *belief* is merely definitional, and has no bearing on the questions about development and cognitive architecture that concern us.[9] Even if infants lack the concept *belief*, it may be that the concept they do use (represented by THINKS, say) develops into a concept of belief through incremental learning. No radical conceptual change, and no introduction of new conceptual primitives, need occur. Moreover, it may be that the conceptual primitives used in the early-developing system are gradually enriched through learning, in such a way that the early-developing system itself *becomes* the one-and-only adult system over developmental time. So even if infants lack the adult conception of belief, this is fully consistent with a one-system account.

It may be that Butterfill and Apperly (2013) are tacitly committed to some version of inferential-role semantics (as are most other cognitive scientists; see Block 1986; Carey 2009). That is, they may think that it is the inferences that people are capable of that partly determine the concepts they possess. And on a holistic construal of inferential-role theories of conceptual content, of course, *any* change in the inferential liaisons of a given representation will lead to the existence of a novel concept. In that case it will be trivially true that adult and infant concepts of belief are distinct. But nothing of any developmental interest would follow from this, unless it could also be shown that the process is not one of continuous inferential enrichment of a single set of representational primitives, but rather involves the creation of a novel set, while leaving the infants' representations to continue operating untouched in adults.[10] The mere fact of conceptual differences between infancy and adulthood would provide no evidence of these further claims. And the definitional question of where, in the process of incremental learning, the concept of *belief* first emerges is of no scientific interest.

Consider, by way of comparison, what one should say about the acquisition of a concept like *cat*. Imagine an infant who has only just mastered the categorical distinctions among cats, dogs, squirrels, and other things. But at this early stage, we may suppose, the infant does not know that cats have kittens, that cats die, that cats need to eat and excrete, that cats feed their young on milk, and so on. This might lead many to say that the infant does not yet possess the adult concept of *cat*. But no one would think, plainly, that this gives us reason to believe that people have two distinct systems for thinking about cats—an early developing one, which enables infants to track the presence of cats without representing them as such, and a later developing one that contains the mature concept. On the contrary, we surely think that the infant concept of *cat* transitions gradually into the adult one, as the infant first learns one fact about cats and then another. So it is, I suggest, with the concept of *belief*.

Let me provide an illustration of the sort of incremental learning that could result in infants understanding the aspectuality of belief for the first time.[11] Suppose that infants have to learn (perhaps from their own experience) that hearing gives rise to thoughts about hearable properties and that sight gives rise to thoughts about visible properties, but not vice versa. And suppose they also learn about the effects of blindfolds on sight

[9] This is one of the places where philosophers, with their traditional focus on definitions and conceptual analysis, have had a bad influence on the rest of cognitive science.

[10] Note that this would require the infant system to be encapsulated from the rest of cognition, in such a way that new inferential liaisons cannot be added, just as Butterfill and Apperly (2013) claim.

[11] The proposals contained in this paragraph seem ripe for experimental testing. Note that they make minimal executive demands.

Author's personal copy

(again from their own experience.) This should then enable them to appreciate that different aspects of the same event can be known without knowing others. For example, with eyes covered, an agent might show a preference for playing with a red rattling ball (having tried out, but shown no interest in, a blue tinkling one). When the agent is then presented with a range of different colored balls with eyes uncovered, infants might be surprised if the agent were to reach immediately for the red ball. They should expect that the agent would first need to test the various balls to see whether or not they make the desired rattling sound.

Note that what has to be learned here are generalizations about what kinds of perceptual access will give rise to what sorts of propositional thought. In the absence of sight, the agent will think THAT BALL IN MY HAND RATTLES, whereas in the presence of sight the agent can see that the rattling ball is red, and hence think THE *RED* BALL RATTLES. Distinct propositions are attributed as the content of belief in the two cases. But no radical conceptual change is required for an infant to become capable of distinguishing between the two kinds of thought-attribution. What is needed is just an increase in the sophistication of the rules for attributing thoughts to others.[12]

I conclude, then, that not only is there evidence that infants can represent the aspectuality of belief (as discussed in Section 2.1), but even if it were true that they cannot do so, this would provide no support for a two-systems account of mindreading ability.

2.3 The Normativity of Belief

Butterfill and Apperly (2013) also argue that the early-developing system cannot represent beliefs as such because beliefs are essentially *normative*—it is said to be part of the concept of *belief* that there are facts about what someone *should* or *should not* believe given the evidence—whereas infants show no sensitivity to such normative facts. This is a disputable claim about our (adult) concept of belief (Glüer and Wikforss 2009). Of course there *are* truths about what people should or should not believe, but it is contentious whether or not someone needs to know these truths in order to grasp the concept of *belief*. But this dispute is of no interest for the developmental debates, unless it can be shown that acquisition of the adult concept requires radical (non-conservative) conceptual change, or unless it can be shown that the infant concept that tracks belief remains operative alongside the as-such concept deployed by adults. No convincing evidence has been presented for either of these claims at this point.

[12] Note that this perspective is really quite close to the "minimal mindreading" ideas of Butterfill and Apperly (2013), with the difference that they disallow any role for learning in the early-developing system, and with the difference that this system is stipulated to be incapable to tracking identities or quantified states of affairs. The former stipulation makes sense if the proposal is that the system should continue to exist unchanged into adulthood. But the latter is odd. If infants themselves can represent identity facts and quantified facts, why should they be prevented from attributing mental states to others that track such facts? If an infant is shown that all the balls in a box are red, for example, and can judge that they are, and then another agent is seen to look inside the box, why should the infant be incapable of judging, THE AGENT REGISTERS: ALL THE BALLS IN THE BOX ARE RED? Of course it might be said that infants are incapable of thinking propositional thoughts at all. But this would be a bold claim. (Remember that propositions don't need to have language-like structure. They just need to encode properties and relations among individuals or collections of individuals.) And it is one for which we have been offered no evidence.

Author's personal copy

Consider children's acquisition of the concept *cat* once again. One can dispute what it is that someone needs to know in order to grasp the adult concept of *cat*. But however that dispute turns out, if any support is to be provided for a two-systems view of people's thinking about cats, then we need to be shown both that some kind of radical (non-conservative) conceptual change takes place and that the early-developing concept continues to exist alongside the adult one. The mere fact of conceptual change provides no evidence either way.

Apperly and Butterfill (2009) draw on a different comparison. They say that the early-developing mindreading system is like the approximate numerosity system that is known to be present in young infants (Izard et al. 2009), and which continues to exist alongside the adult exact number system that children only acquire much later (Barth et al. 2003). But there are two reasons why we should be suspicious of this analogy. One is that the exact number system is only acquired laboriously by children as a result of explicit training with the count list and with counting procedures (Carey 2009). In contrast, no explicit teaching of mindreading skills takes place (which isn't to say that no learning happens as a result of infants being exposed to talk about mental states, of course). And the second reason is that systems of exact number are not a human universal (Pica et al. 2004), whereas a capacity to reason explicitly about false beliefs *is* universal (Callaghan et al. 2005).[13] This suggests that exact number concepts are culturally constructed in a way that core mindreading concepts are not.

2.4 The Complexity of Belief

The final theoretical argument in support of their two-systems view that Butterfill and Apperly (2013) present concerns the *complexity* of propositional attitudes like belief. On any account, propositional attitudes possess complex causal structures. For one thing, they allow for arbitrarily complex nestable propositional contents. Jane can believe that *Mary is tall*, or she can believe that *Sue thinks Mary is tall*, or she can believe that *Fred is sad that Sue thinks Mary is tall*, and so on, indefinitely. Moreover, propositional attitudes enter into complex interactions with other such attitudes. The causal chains of thought that can issue in a novel belief, decision, or action can be arbitrarily long and convoluted. This complexity will need to be mirrored, Butterfill & Apperly think, in the complexity of the concepts that are required for representing propositional attitudes as such. As a result it is said to be unlikely that new judgments involving such concepts could be formed swiftly or automatically. In contrast, as we will see in Section 3, other forms of mindreading *can* be swift and automatic, suggesting that the concepts involved in each are distinct. And it is equally unlikely, it is said, that young infants should be capable of such complex judgments, suggesting that they rely on the same concepts that are involved in swift and automatic mindreading.

The complexity of the concepts appealed to here is of two sorts, however. One derives from the possibility of recursive embedding. But a mere *possibility* cannot make simple (non-iterated) attributions cognitively demanding. It may well be that thoughts with multiply-embedded contents can neither be grasped by infants nor

[13] Note that one of the authors of this study is Lillard, who had previously (1998) suggested that false-belief reasoning is *not* a human universal.

Author's personal copy

processed swiftly and automatically. But this might be because of the processing resources that are required to construct and entertain such thoughts, rather than any inherent limitation in the systems or concepts involved. Consistent with these suggestions, McKinnon and Moscovitch (2007) find that concurrent working-memory load has little impact on adult performance in first-order false-belief tasks, while having a big impact on performance in second-order tasks involving multiply-embedded contents.

A second source of complexity derives from the inferential liaisons into which propositional-attitude concepts can enter. These may, indeed, be arbitrarily complex. But there is no reason to think that all of this complexity should be activated and explicitly attributed in any given case. It remains possible that representations of propositional attitudes are sometimes tokened swiftly and automatically (given appropriate allocations of attention), even if reasoning about some of the more sophisticated inferences into which those representations can enter would be cognitively demanding.

Butterfill and Apperly (2013) seem to think that there is something especially cognitively demanding about the attribution of *propositional* contents as such, hence placing such attributions beyond the reach of infants. But it is quite unclear why this should be so. For if we assume, as we surely should, that infants are capable of propositional thoughts such as THE RED BALL IS IN THE BOX, and if we allow that infants at least possess a concept like REGISTERS, then all the infant needs to do is embed that very representation into the scope of an attribution, thinking, THE AGENT REGISTERS: THE RED BALL IS IN THE BOX. The result is that the infant has attributed a propositional content to the agent.[14]

Butterfill & Apperly may retort that this is not yet to attribute propositional contents *as such*, because (they might insist) infants lack the concept of a *proposition*. But it is likely that many older children and adults likewise lack such a concept. The difference, of course, is that adults will (whereas infants may not; but see Section 2.1) modulate their attribution of one proposition rather than another (extensionally equivalent) one, depending on the epistemic situation of the agent. Only older children and adults might thus be said to possess at least an *implicit* concept of a proposition. But notice that there is no reason to think that one cannot get from the infant concept to the adult one via incremental learning and conceptual change, as we saw in Section 2.2. So there is no reason, here, to postulate the existence of two distinct systems.

I conclude, then, that none of the theoretical considerations adduced by Apperly and Butterfill (2009) and Butterfill and Apperly (2013) in support of the two-systems hypothesis is capable of carrying much weight. (Nor is any of the empirical evidence discussed so far.) If there is a case to be made, it must derive largely from experimental evidence of two systems in adults. This will form the topic of Section 3. (For more extensive discussion of the adult data than there is space for here, see Carruthers 2015a.)

[14] Perhaps Butterfill & Apperly believe that infants are incapable of thinking propositional thoughts at all. This would then explain why they think infants are incapable of attributing propositional thoughts to other people. But this would be tantamount to claiming that infants are incapable of *thinking*. For although thoughts can be described and individuated either extensionally or intensionally for various purposes, thoughts themselves are inherently aspectual (and hence propositional) in nature. Anything that an infant can think will represent some aspects of a state of affairs but not others.

Author's personal copy

3 Fast and Automatic Versus Slow and Effortful

Butterfill and Apperly (2013) appeal to data collected with adults to suggest that visual perspectives are sometimes computed swiftly and automatically, but sometimes in a way that is slow and effortful. Similarly, they appeal to other evidence suggesting that false beliefs can be tracked swiftly and automatically, but also slowly and effortfully. Both forms of evidence will be discussed below. Butterfill & Apperly reason that since the same system cannot be both fast and automatic *and* slow and effortful, there must be two distinct systems at work. In a weak sense, this is correct: reasoning intuitively is different from reasoning reflectively. But as we will see, it does not follow that distinct conceptual resources are involved.

3.1 Representing Visual Perspectives

Samson et al. (2010) show that people cannot easily ignore what other people can see. When judging what they themselves can see in a display containing an avatar, they are slower and make more errors in circumstances where there is a conflict between what the avatar can see and what they can see. This suggests that the perceptual state of the avatar has been computed swiftly and automatically, interfering with the participants' responses when reporting what they themselves can see. Qureshi et al. (2010) show in addition that these computations of others' perspectives are not impacted by cognitive load, further supporting the claim that the computations are fully automatic, and are independent of executive resources.

In these experiments the perceptual judgments are quite simple ones. Participants judge whether they can see (or in some conditions, whether the avatar can see) one dot or two. Judgments of this sort could be handled by the early-developing system postulated to be present in infancy, which computes whether an agent *registers* something, but without being capable of computing the *manner in which* the thing is registered by the agent, or the *aspect* under which the thing gets perceived. In support of this interpretation, Surtees et al. (2012) find that aspectual judgments are *not* computed automatically. In conditions where participants have to report whether they see a "6" or a "9", they are no slower and no less accurate when the avatar can see the same figure differently, resulting from his distinct visual perspective on the stimulus. (Note that someone seeing a "6" upside-down will see it as a "9".)

Taken together, Butterfill and Apperly (2013) claim that these two bodies of evidence support their two-systems account. But they are equally consistent with a one-system view, since the two tasks differ in their executive demands. In order to judge whether someone can see one dot or two one just needs to compute lines of sight (judging whether or not one of the dots falls outside the person's field of vision). But in order to judge whether someone sees a "6" or a "9", in a case where one sees the stimulus as a "6" oneself, one needs to take one's own image of the stimulus and mentally rotate it to an extent appropriate for the person's spatial position and orientation. This is executively demanding, and makes constitutive use of working memory to sustain and manipulate one's visual representation. But the conceptual resources involved in the two tasks need not differ at all. Both can involve full-blown judgments of *seeing*.

Author's personal copy

Indeed, data from the same lab actually support this one-system / executive-function reading. Surtees et al. (2013) find that visual rotation of the stimulus is required *both* for judgments of how something visually appears to a target agent *and* for judgments of whether something is positioned to the target's left or right. In contrast, neither judgments of whether something is in front of or behind the target, nor simple judgments of whether the target can see something or not, require mental rotation. Hence the same mental-rotation explanation should be advanced of both the difference between judgments of left / right and judgments of front / behind, on the one hand, and of the difference between judgments of see / not-see and judgments of sees-6 / sees-9, on the other. The latter contrast should not, then, be explained by postulating the existence of two different mindreading systems with distinct conceptual resources. Rather, we can suppose that there is one such system that either does, or does not, operate together with executively-controlled mental manipulations of representations of the stimuli, depending on the task demands.

3.2 Representing Beliefs

Kovács et al. (2010) provide evidence that adults automatically compute what another agent will falsely believe about a display (or at least, falsely *register*), with the resulting information interfering with the participants' own expectations about the situation. In this experiment it seems that people tacitly put as much credence in the beliefs of others as they do in their own beliefs. van der Wel et al. (2014) replicate the basic finding that people automatically compute and represent the false beliefs of others, and they find that the trajectory of people's reaching actions are influenced by those beliefs. But in this case they find that others' beliefs about the location of a target have *lesser* influence on one's actions than do one's own beliefs.

In contrast with these findings, Back and Apperly (2010) provide evidence that adults do *not* automatically represent what another agent falsely believes, since they are slower to make reports of people's beliefs than they are to report the location of the target object. Again this is said to support a two-systems theory. Butterfill and Apperly (2013) claim that it is the early-developing system that is responsible for the positive results in Kovács et al. (2010) (and presumably also those of van der Wel et al. 2014). But by hypothesis, that system cannot inform verbal reports. Hence the findings of Back and Apperly (2010) can be taken as evidence that the system that guides verbal report does not operate automatically. However, again these findings admit of an alternative one-system / executive-function explanation.

In the experiments conducted by Kovács et al. (2010), participants watched while an object moved behind a screen, emerged, and either moved off-stage or returned behind the screen. (van der Wel et al. 2014, used a similar display; but their dependent measure was a reach toward the target object rather than a simple button-press.) Meanwhile an agent was either present or absent in the display, and saw some or all of the proceedings while playing no active part in them. The participants' task was simply to press a button as fast as possible if the object was present when the screen dropped flat at the end of the sequence. Naturally, reaction times were faster when participants expected the object to be there. But they were also faster when the incidental agent expected the object to be there, even though the participants themselves had seen it leave the stage. To explain the data in this experiment we need only suppose that the agent's belief is

Author's personal copy

computed and stored, thereby priming the appropriate action (pressing the button, in Kovács et al.'s experiment; or priming a direction of reach in the experiments of van der Wel et al.). No executive resources need to be appealed to in our explanation of the effects.

In the experiments conducted by Back and Apperly (2010), in contrast, participants watched while a man played a sort of "shell game" with an object and two cups, placing the object now under one cup, now under another. The participants' overt task was to answer a question about the location of the object when probed at some point during the sequence. Another agent watched some or all of each episode, being present for some of the time, while leaving the room at others. On unexpected belief-probe trials participants were asked where the second agent believed the object to be. They were slower and made more errors when answering (in both the true-belief and false-belief conditions) than they were when answering the location-of-the-object question. (In contrast, when participants were asked to keep track of the agent's beliefs as well as the location of the object, there were no differences in reaction times and error rates. So the two tasks were equally difficult.) This is taken to show that the second agent's beliefs about location had not *already* been computed automatically as the sequence unfolded, but needed to be calculated retrospectively when probed (hence the lag in time and the increase in errors).

The data can equally well be explained, however, in terms of the executive cost of accessing a stored memory to serve as the basis for a report. It may be that the incidental agent's beliefs *were* automatically computed and stored. But since they were seemingly irrelevant to the task requirements these representations weren't held active in working memory. When subjects are unexpectedly probed, then, they need to access and recall the stored beliefs formed earlier. This will take time, of course, and we can suppose that memory-search is error-prone. So the data can be fully explained consistent with a one-system account of mindreading, maintaining that there is just a single set of conceptual resources involved.[15]

This interpretation can be further strengthened by considering the findings of Schneider et al. (2012b). They build on the results of Schneider et al. (2012a), who use eye-tracking measures to show that people unconsciously compute representations of the true and false beliefs of other agents (or at least true and false registerings), and form expectations on their basis, while engaged in an unrelated task. But Schneider et al. (2012b) show that these expectations are absent when participants are placed under cognitive load. In such circumstances their eye movements no longer reveal mental-state understanding. From a two-systems perspective these findings suggest that even the early-developing system doesn't operate automatically, but depends rather on executive resources. (Note that anticipatory looking is thought to be driven by the output of the early-developing system.) But that would be in tension with the findings of Qureshi et al. (2010), which suggest that the early-developing system does *not* depend on executive load. In fact both sets of findings may be better explained by a

[15] And just as this account predicts, if the delay between participants observing the belief-inducing event and the unexpected belief-probe is greatly reduced, people do *not* show any cost in reaction times or error rates (Cohen and German 2009). This is presumably because the belief-representation is still readily available, and does not need to be retrieved from long-term memory.

Author's personal copy

one-system account, once appropriate consideration is given to the varying task demands.

As already noted, in the tasks used by Qureshi et al. (2010) participants merely had to compute the line of sight of the avatar involved, with the results of these computations priming or counter-priming the participants' own reactions. No executive resources were required. In the tasks used by Schneider et al. (2012a, 2012b, 2014a), in contrast, participants needed to follow the course of a visually-presented false-belief narrative that took a full minute to unfold, but which was irrelevant to their goals. (Their task was to press the space-bar as swiftly as possible whenever the actor in the narrative waved at the camera.) In order to form expectations of the agent's behavior, her beliefs about the location of the object would first need to be computed. (Then later when the object is moved in her absence, this representation would need to be left untouched.) That representation of the agent's belief would then either need to be sustained in working memory while the story unfolds, or it would need to be retrieved from long-term memory at the end of the sequence.

It seems from Schneider et al. (2012a, 2014a) that these processes take place unconsciously and independently any overt goal, since participants show no awareness of representing the agent's beliefs even when probed at length afterwards, and since the findings aren't modulated by task instructions. Hence if one accepts that working-memory contents are always conscious (Prinz 2012; Carruthers 2015b), we can conclude that a representation of the agent's belief would need to have been retrieved from long-term memory at the appropriate point in the story (when the agent's reaching action is cued by the sound of a bell). If this analysis is correct, then Schneider et al.'s (2012b) finding that there is no behavioral evidence of belief-tracking when participants are placed under executive load is just what one might expect. For initiating a memory-search and subsequent retrieval of memories surely requires executive resources. These data are thus consistent with a one-system account according to which computations of belief are automatic (given appropriate allocations of attention), whereas the later use of those representations may depend on executive resources.

3.3 Two Systems Revisited

I have suggested that the data collected with adults that is alleged to support a two-systems account may be better explained by the distinction between tasks that do, or do not, make constitutive use of the resources of working memory, together with those that do, or do not, require the use of mnemonic cues to retrieve previously processed mindreading information. It can be just one mindreading system, with one set of conceptual resources, that is involved throughout. The resulting theory *can* be conceptualized as a kind of two-systems account, of course, one of which is automatic and the other of which depends on executive function. But in the present case one system encompasses the other, with the automatic system forming a proper part of the executively controlled one. And the very same conceptual resources are deployed by each. So this is a two-systems theory of a very different sort.

Author's personal copy

4 Conclusion

I have argued that the two-systems framework advanced by Apperly and colleagues is neither supported by theoretical considerations, nor is it adequately supported by the existing evidence. If I am right, future tests of the two-systems account will need to be much more careful in matching the executive and mindreading demands of the tasks employed. I have also argued that both theory and data suggest that the infant mindreading system develops gradually, transforming into the adult one through incremental learning and piecemeal conceptual change, and does not continue to exist alongside the latter. Tasks that implicate this system will differ, however, in the extent to which they require the mindreading system to operate in concert with executively-controlled working-memory processes and/or targeted searches of long-term memory. This proposal issues in a two-systems account of a different and much weaker sort. The resulting account is consistent with the suggestion of incremental conceptual change, and maintains that the same set of conceptual resources is drawn on in all of the adult tasks.

Acknowledgments I am grateful to Suilin Lavelle, Evan Westra, and two anonymous referees for their comments on an earlier draft of this paper.

References

Apperly, I. 2011. *Mindreading*. New York: Psychology Press.

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

Apperly, I., K. Riggs, A. Simpson, C. Chiavarino, and D. Samson. 2006a. Is belief reasoning automatic? *Psychological Science* 17: 841–844.

Apperly, I., D. Samson, N. Carroll, S. Hussain, and G. Humphreys. 2006b. Intact first- and second-order false belief reasoning in a patient with severely impaired grammar. *Social Neuroscience* 1: 334–348.

Astington, J., and J. Jenkins. 1999. A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology* 35: 1311–1320.

Back, E., and I. Apperly. 2010. Two sources of evidence on the non-automaticity of true and false belief ascription. *Cognition* 115: 54–70.

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

Baillargeon, R., Z. He, P. Setoh, R. Scott, S. Sloan, and D. Yang. 2013. False-belief understanding and why it matters: the social-acting hypothesis. In *Navigating the social world*, ed. M. Banaji and S. Gelman. Oxford: Oxford University Press.

Barrett, H., Broesch, T., Scott, R., He, Z., Baillargeon, R., Wu, D., Bolz, M., Henrich, J., Setoh, P., Wang, J., and Laurence, S. 2013: Early false-belief understanding in traditional non-western societies. *Proceedings of the Royal Society, B (Biological Sciences)*, 280, #1755.

Barth, H., N. Kanwisher, and E. Spelke. 2003. The construction of large number representations in adults. *Cognition* 86: 201–221.

Block, N. 1986: An advertisement for a semantics for psychology. In P. French, T. Euhling, and H. Wettstein (eds.), *Midwest Studies in Philosophy: 10: Studies in the Philosophy of Mind*, University of Minnesota Press.

Buttelmann, D., M. Carpenter, and M. Tomasello. 2009. Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112: 337–342.

Buttelmann, D., H. Over, M. Carpenter, and M. Tomasello. 2014. Eighteen-month-olds understand false beliefs in an unexpected-contents task. *Journal of Experimental Child Psychology* 119: 120–126.

Author's personal copy

Buttelmann, F., J. Suhrke, and D. Buttelman. 2015. What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology* 131: 94–103.

Butterfill, S., and I. Apperly. 2013. How to construct a minimal theory of mind. *Mind & Language* 28: 606–637.

Call, J., and M. Tomasello. 1999. A nonverbal false belief task: the performance of children and great apes. *Child Development* 70: 381–395.

Callaghan, T., P. Rochat, A. Lillard, M. Claux, H. Odden, S. Itakura, S. Tapanya, and S. Singh. 2005. Synchrony in the onset of mental-state reasoning. *Psychological Science* 16: 378–384.

Carey, S. 2009. *The origin of concepts*. New York: Oxford University Press.

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

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

Carruthers, P. 2013. Mindreading in infancy. *Mind & Language* 28: 141–172.

Carruthers, P. 2015a: Mindreading in adults: Evaluating two-systems views. *Synthese*, 196.

Carruthers, P. 2015b. *The centered mind: what the science of working memory shows us about the nature of human thought*. New York: Oxford University Press.

Clements, W., and J. Perner. 1994. Implicit understanding of belief. *Cognitive Development* 9: 377–397.

Cohen, A., and T. German. 2009. Encoding of others' beliefs without overt instruction. *Cognition* 111: 356–363.

De Bruin, L., and A. Newen. 2012. An association account of false belief understanding. *Cognition* 123: 240–259.

Gallagher, S., and D. Povinelli. 2012. Enactive and behavioral abstraction accounts of social understanding in chimpanzees, infants, and adults. *Review of Philosophy and Psychology* 3: 145–169.

Glüer, K., and Å. Wikforss. 2009. Against content normativity. *Mind* 118: 31–70.

Gopnik, A., and A. Meltzoff. 1997. *Words, thoughts, and theories*. Cambridge: MIT Press.

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

Izard, V., C. Sann, E. Spelke, and A. Streri. 2009. Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences* 106: 10382–10385.

Kloo, D., and J. Perner. 2003. Training transfer between card sorting and false belief understanding: helping children apply conflicting descriptions. *Child Development* 74: 1823–1839.

Knudsen, B., and U. Liszkowski. 2012. 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy* 17: 672–691.

Kovács, Á. 2009. Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science* 12: 48–54.

Kovács, Á., E. Téglás, and A. Endress. 2010. The social sense: susceptibility to others' beliefs in human infants and adults. *Science* 330: 1830–1834.

Leslie, A., O. Friedman, and T. German. 2004. Core mechanisms in "theory of mind. *Trends in Cognitive Sciences* 8: 528–533.

Lillard, A. 1998. Ethnopsychologies: cultural variations in theories of mind. *Psychological Bulletin* 123: 3–32.

Low, J., and J. Watts. 2013. Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science* 24: 305–311.

Low, J., W. Drummond, A. Walmsley, and B. Wang. 2014. Representing how rabbits quack and competitors act: limits on preschooler's efficient ability to track perspective. *Child Development* 85: 1519–1534.

Luo, Y. 2011. Do 10-month-old infants understand others' false beliefs? *Cognition* 121: 289–298.

Luo, Y., and W. Beck. 2010. Do you see what I see? Infants' reasoning about others' incomplete perceptions. *Developmental Science* 13: 134–142.

McKinnon, M., and M. Moscovitch. 2007. Domain-general contributions to social reasoning: theory of mind and deontic reasoning re-explored. *Cognition* 102: 179–218.

Milligan, K., J. Astington, and L. Dack. 2007. Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development* 78: 622–646.

Moll, H., and A. Meltzoff. 2011. How does it look? Level 2 perspective-taking at 36 months of age. *Child Development* 82: 661–673.

Onishi, K., and R. Baillargeon. 2005. Do 15-month-olds understand false beliefs? *Science* 308: 255–258.

Author's personal copy

Perner, J. 2010. Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In *Cognition and neuropsychology: international perspectives on psychological science: volume 1*, ed. P. Frensch and R. Schwarzer. New York: Psychology Press.

Perner, J., and T. Ruffman. 2005. Infants' insight into the mind: how deep? *Science* 308: 214–216.

Perner, J., T. Ruffman, and S. Leekam. 1994. Theory of mind is contagious: you catch it from your sibs. *Child Development* 65: 1228–1238.

Peterson, C., and M. Siegal. 1995. Deafness, conversation and theory of mind. *Journal of Child Psychology and Psychiatry* 36: 459–474.

Pica, P., C. Lemer, V. Izard, and S. Dehaene. 2004. Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306: 499–503.

Poulin-Dubois, D., and V. Chow. 2009. The effect of a looker's past reliability on infants' reasoning about beliefs. *Developmental Psychology* 45: 1576–1582.

Prinz, J. 2012. *The conscious brain*. Oxford: Oxford University Press.

Qureshi, A., I. Apperly, and D. Samson. 2010. Executive function is necessary for perspective-selection, not level-1 visual perspective-calculation: evidence from a dual-task study of adults. *Cognition* 117: 230–236.

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

Samson, D., I. Apperly, J. Braithwaite, B. Andrews, and S. Bodley Scott. 2010. Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance* 36: 1255–1266.

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

Schneider, D., A. Bayliss, S. Becker, and P. Dux. 2012a. Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General* 141: 433–438.

Schneider, D., R. Lam, A. Bayliss, and P. Dux. 2012b. Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science* 23: 842–847.

Schneider, D., Z. Nott, and P. Dux. 2014a. Task instructions and implicit theory of mind. *Cognition* 133: 43–47.

Schneider, D., V. Slaughter, S. Becker, and P. Dux. 2014b. Implicit false-belief processing in the human brain. *NeuroImage* 101: 268–275.

Scott, R., and R. Baillargeon. 2009. Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development* 80: 1172–1196.

Scott, R., R. Baillargeon, H. Song, and A. Leslie. 2010. Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology* 61: 366–395.

Scott, R., Z. He, R. Baillargeon, and D. Cummins. 2012. False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous-response tasks. *Developmental Science* 15: 181–193.

Scott, R., Richman, J., and Baillargeon, R. 2015: Infants understand deceptive intentions to implant false beliefs: New evidence for early mentalistic reasoning. *Cognitive Psychology*, 78

Senju, A., V. Southgate, C. Snape, M. Leonard, and G. Csibra. 2011. Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science* 22: 878–880.

Song, H., and R. Baillargeon. 2008. Infants' reasoning about others' false perceptions. *Developmental Psychology* 44: 1789–1795.

Song, H., K. Onishi, R. Baillargeon, and C. Fisher. 2008. Can an actor's false belief be corrected by an appropriate communication? Psychological reasoning in 18.5-month-old infants. *Cognition* 109: 295–315.

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

Southgate, V., A. Senju, and G. Csibra. 2007. Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science* 18: 587–592.

Southgate, V., C. Chevallier, and G. Csibra. 2010. Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science* 13: 907–912.

Surian, L., S. Caldi, and D. Sperber. 2007. Attribution of beliefs by 13-month-old infants. *Psychological Science* 18: 580–586.

Surtees, A., and I. Apperly. 2012. Egocentrism and automatic perspective-taking in children and adults. *Child Development* 83: 452–460.

Surtees, A., S. Butterfill, and I. Apperly. 2012. Direct and indirect measures of Level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology* 30: 75–86.

Surtees, A., I. Apperly, and D. Samson. 2013. Similarities and differences in visual and spatial perspective-taking processes. *Cognition* 129: 426–438.

Author's personal copy

Träuble, B., V. Marinovic, and S. Pauen. 2010. Early theory of mind competencies: Do infants understand others' beliefs? *Infancy* 15: 434–444.

van der Wel, R., N. Sebanz, and G. Knoblich. 2014. Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition* 130: 128–133.

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

Woolfe, T., S. Want, and M. Siegal. 2002. Signposts to development: theory of mind in deaf children. *Child Development* 73: 768–778.

Yott, J., and D. Poulin-Dubois. 2012. Breaking the rules: Do infants have a true understanding of false belief? *British Journal of Developmental Psychology* 30: 156–171.