Forum

CellPress

Making sense of early false-belief understanding

Katharina A. Helming[1,2], Brent Strickland[1], and Pierre Jacob[1]

[1] Institut Jean Nicod, Centre National de la Recherche Scientifique, Ecole des Hautes Etudes en Sciences Sociales, Ecole Normale Supérieure, Pavillon Jardin, 29 rue d'Ulm, F-75005 Paris, France
[2] Differential Psychology, Personality Psychology, and Diagnostics, Department I, University of Trier, 54286 Trier, Germany

We address the puzzle about early belief ascription: young children fail elicited-response false-belief tasks, but they demonstrate spontaneous false-belief understanding. Based on recent converging evidence, we articulate a pragmatic framework to solve this puzzle. Young children do understand the contents of others' false belief, but they are overwhelmed when they must simultaneously make sense of two distinct actions: the instrumental action of a mistaken agent and the experimenter's communicative action.

The puzzle

When asked to predict where Sally, who falsely believes her toy to be in a green container, will look for her toy (elicited-response false-belief task), most 3-year-olds who know it to be in a blue container point to the blue container, not the green one [1] (Figure 1). Yet, 15-month-olds have been shown to spontaneously look longer when an agent reaches for her toy either at its actual location while she falsely believes it to be elsewhere or elsewhere while she knows its actual location [2]. Children under 3 years have also been shown to correctly gaze at an empty location where an agent falsely believes her toy to be in anticipation of the agent's action [3].

Why do most 3-year-olds fail elicited-response tasks, whereas much younger children spontaneously either anticipate where a mistaken agent will look for her toy or look longer when an agent fails to act in accordance with what she truly or falsely believes? This is a major puzzle in the study of early social cognition. There are presently two strategies for addressing this puzzle.

The cultural constructivist approach takes success at elicited-response tasks as a necessary condition of the ability to ascribe false beliefs to others, construed as the output of 'a cultural process tied to language acquisition' [4]. If so, failure at elicited-response false-belief tasks demonstrates the inability to ascribe false beliefs to others. Thus, some leading advocates of this approach have recently proposed that young children predict that Sally will look for her toy at its actual location because they construe an agent's action in terms of what makes objective sense for her, not in terms of her subjective mentalistic reasons [5].

By contrast, the processing-load account argues that young children could succeed spontaneous-response false-belief tasks while failing elicited-response tasks for other reasons; for example, if they lack the executive resources required to inhibit the content of their own knowledge of the toy's location and to select the content of the agent's false belief in response to the experimenter's question. The overwhelming demands of the task generate in young children a reality bias [6].

The limits of the processing-load account

However, in a novel false-belief task involving a puppet (the Duplo girl) who had a false belief about the location of her bananas, 3.5-year-olds (who knew the actual location of the bananas) were prompted to act out the puppet's most likely action by being told 'What happens next? You can take the girl yourself if you want. What is she going to do now?' Most 3.5-year-olds moved the girl to the empty location [7].

To take the Duplo girl to the empty location, 3.5-year-olds must have inhibited their own knowledge and selected the content of the girl's false belief. Why did the experimenter's prompt not overwhelm their inhibitory resources and generate a reality bias, as predicted by the processing-load account? The processing-load account clearly needs some 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. We offer a pragmatic explanation.

Perspective taking on instrumental and communicative agency

The full-blown human mind-reading system has evolved to make sense of two kinds of agency: instrumental and communicative. Furthermore, adult mind-readers can take either a third-person detached perspective or a second-person engaged perspective on both kinds of action (Table 1). Much of the evolutionary pressure for the human ability to track the contents of others' false beliefs derives from the demands of communicative agency: verbal human intentional communication is a unique potential source of novel information, but also of possible misinformation [8,9].

Young children spontaneously track the contents of the motivations and epistemic states (including the false beliefs) of agents of instrumental actions from a third-person perspective [2]. They also give evidence of their

Corresponding author: Jacob, P. (piotrjacob@gmail.com).

1364-6613/$ – see front matter
 http://dx.doi.org/10.1016/j.tics.2014.01.005

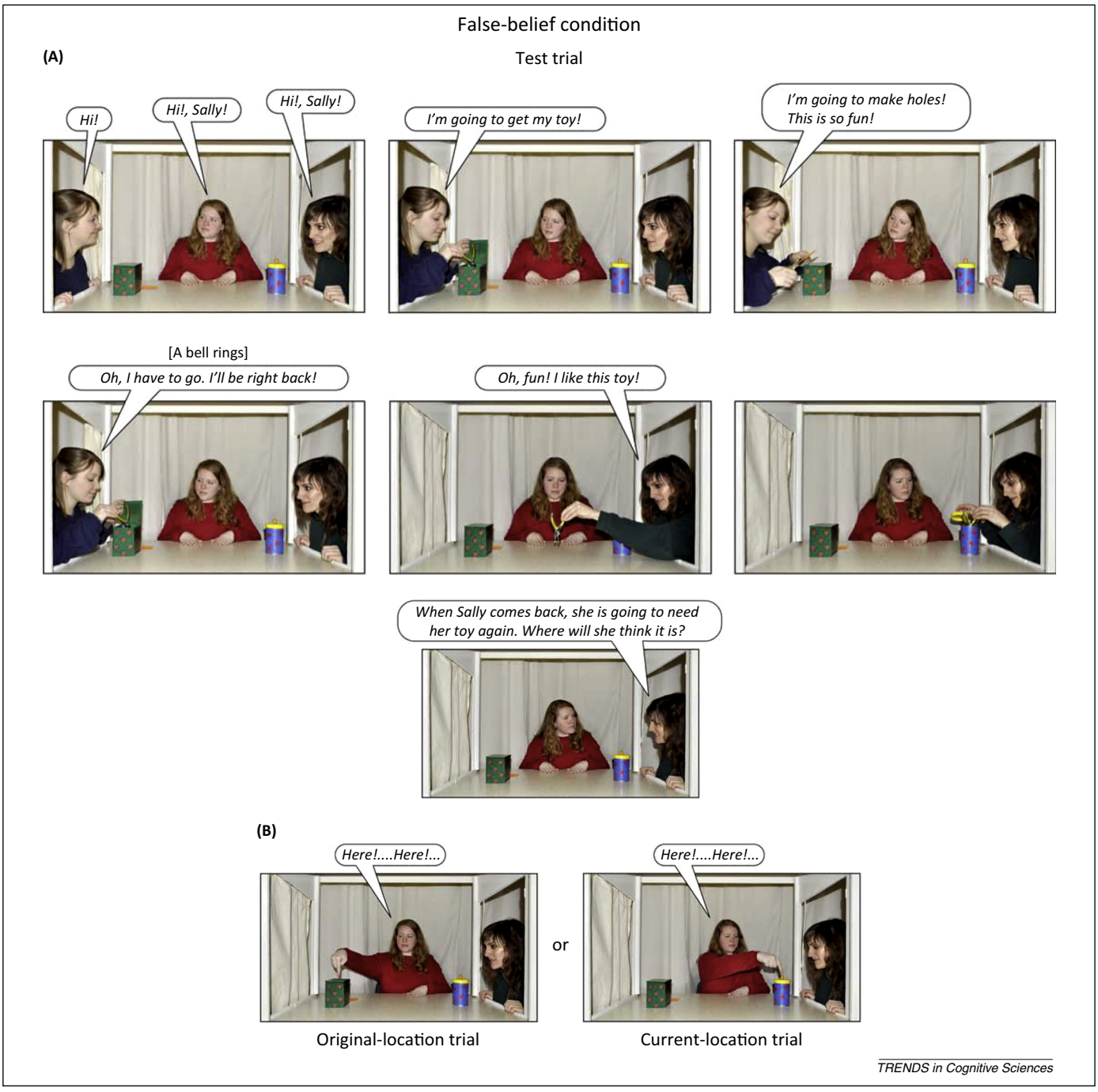

Figure 1. Watching an adult pass a false-belief task. **(A)** After placing her toy in the green box, Sally (left) leaves. While she is away, the experimenter (right) moves the toy from the green to the blue box. Now Sally falsely believes her toy to be in the green box. If asked to predict where Sally will look for her toy, most 3-year-olds point to the blue box [1]. This shows that they cannot simultaneously keep track of the content of Sally's false belief from a third-person perspective and engage with the experimenter's communicative action from a second-person perspective. **(B)** However, if 2.5-year-olds watch while an adult subject (middle) is asked to predict where Sally will look for her toy, they look longer when the adult points to the blue rather than to the green box [14]. This shows that 2.5-year-olds can track the content of Sally's false belief from a third-person perspective when they are not simultaneously requested to engage with the experimenter's communicative action from a second-person perspective.

altruistic motivation to help a mistaken agent achieve her goal by pointing to the target's actual location, thereby taking a cooperative second-person perspective on the agent's instrumental action [10,11]. Seventeen-month-olds are even able to give to a mistaken agent the intended referent of her pointing gesture, not the object at the demonstrated location [12].

To take a third-person perspective on a speaker's communicative action is to represent the content of the speaker's communicative intention; that is, her intention to display her informative intention (i.e., her wish to convey some new information to her addressee; Table 1, cell 3). A speaker's addressee takes a second-person perspective on the speaker's communicative action if and when the addressee engages in the communicative act and fulfills the speaker's informative intention [8] (Table 1, cell 4). The evidence for so-called natural pedagogy shows that very young children are uniquely sensitive to others' communicative intentions

Table 1. Perspective taking on instrumental and communicative agency

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

When tested in spontaneous-response tasks, preverbal infants take a detached third-person perspective on an agent's instrumental action (cell 1); the evidence shows that they expect the agent to act in accordance with the contents of her true and false beliefs [2,3,6].
When they take a detached third-person perspective on an adult who is being asked to predict where a mistaken agent will reach for her toy (cells 1 and 3), they are surprised if the adult wrongly points to the actual location [14].
When they must simultaneously take a detached third-person perspective on an instrumental action and also engage with the experimenter's communicative action from a second-person perspective (cells 1 and 4), their ability to predict a mistaken agent's likely instrumental action breaks down [1,15].

and that they tend to fulfill others' informative intention by interpreting their nonverbal communicative actions as teaching demonstrations [13].

A pragmatic framework

Young children have been shown to spontaneously track the contents of others' false beliefs [2,3] and to be sensitive to others' communicative intentions [13]. In a nutshell, what makes the standard where-prediction question taxing for young children is that it simultaneously requires them to take a detached (or noncooperative) third-person perspective on the mistaken agent's instrumental action while taking a second-person cooperative perspective on the experimenter's communicative action (Table 1, cells 1 and 4). In support of this view, 2.5-year-olds have been shown to look longer when they see an adult point to the toy's actual location rather than the empty location, if and when the adult is being asked where mistaken Sally will look for her toy in front of them [14] (Figure 2). However, their ability to maintain a third-person perspective and to gaze in anticipation of a mistaken agent's action breaks down when the experimenter addresses them directly, as opposed to thinking out loud [15] (Figure 2 and Table 1).

How does children's second-person engagement with the experimenter's communicative action disrupt their ability to keep track of the content of the instrumental agent's false belief? In answering the question 'Where will Sally look for her toy?' participants have the option of mentally representing either the toy's actual location or the location where Sally mistakenly believes it to be. The experimenter's wording of the question may bias children toward the actual location, by virtue of the fact that, in asking the question, the experimenter refers to the toy while she shares the children's correct perspective on its actual location, at the expense of Sally's incorrect perspective (referential bias).

Furthermore, very young children might feel impelled by their altruistic propensities to help an agent achieve her goal-directed action whose success is being compromised by a false belief caused by someone else (cooperative bias). To help a mistaken agent achieve her goal-directed action is to take a second-person perspective on the agent's action. To be helpful, they can rely on their own true knowledge and point to the object's actual location, in accordance with the reality bias [10–12]. If so, they might turn the experimenter's prediction question into the normative question

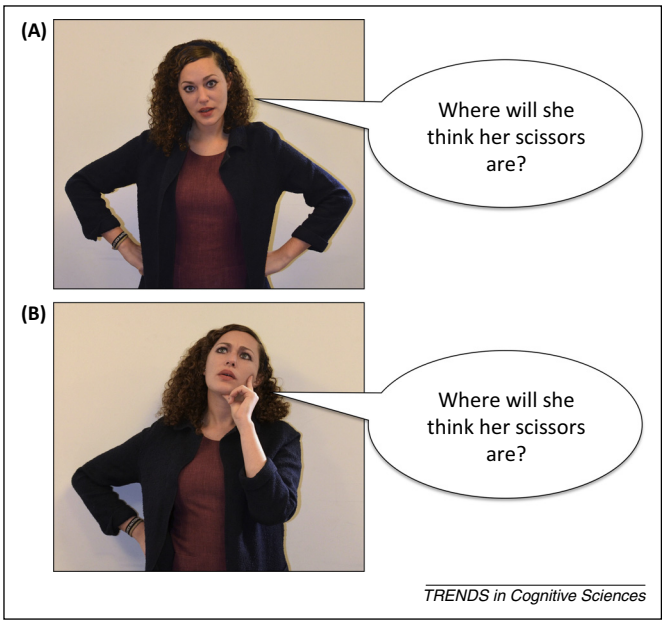

Figure 2. Being addressed or not. Two-and-a-half-year-olds have been shown to gaze correctly at the location where the mistaken agent is likely to reach for her toy in anticipation of the agent's action [3]. This shows that they can predict the action of a mistaken agent by taking a third-person perspective on her action. **(A)** If, however, their third-person perspective on the mistaken agent's action is disrupted by the communicative action of an experimenter who asks them where the mistaken agent will look for her toy, their ability to anticipate the mistaken agent's action breaks down [15]. **(B)** However, if the experimenter makes it clear that the same interrogative sentence is addressed not to the children but to herself, the ability of 2.5-year-olds to anticipate the mistaken agent's action is shown not to break down [15].

'Where *should* Sally look for her toy?' If they do, the correct answer to the normative question is the toy's actual location, not the location where Sally believes it to be.

Concluding remarks

The core insight underlying the pragmatic framework is that a main (if not *the* main) evolved function of the human ability to track the contents of others' false beliefs is to enable humans to deal with false beliefs in the context of communicative agency [8]. Preverbal infants can track the contents of others' false beliefs, but unlike mature speakers they cannot yet intentionally cause others to acquire false beliefs, let alone transmit false beliefs to others via verbal communication. Also unlike mature addressees, they cannot infer true conclusions from misinformation conveyed by speakers. In accordance with this core insight, young children have been shown to become able to deal with others' false testimony approximately when they can pass elicited-response false-belief tasks [9]. So far, the evidence shows that younger children (who fail elicited-response false-belief tests) are prone to help a mistaken agent achieve the goal of her instrumental action [10–12]. There is some scant evidence that being actively involved in causing another's false belief helps improve young children's performance in elicited-response false-belief tasks [1]. This topic urgently needs further detailed investigation. Arguably, 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 by the experimenter into deceiving the girl [7].

The pragmatic framework highlights two biases generated by the experimenter's communicative action. The cooperative-bias hypothesis predicts that children should perform better at elicited-response tasks if they feel less inclined to help the mistaken agent; if, for example, the mistaken agent is an out-group rather than an in-group member. The referential bias rests on two parameters: one is that, by asking the prediction question, the experimenter refers to Sally's toy and thereby draws attention to the toy's actual location. When 3-year-olds are asked a different open question, they take the Duplo girl to the empty location [7]. The second parameter of the referential bias is that the experimenter shares the child's correct epistemic perspective on the actual location of Sally's toy. This parameter could be further tested if the experimenter who asks children to predict Sally's action would share either the children's correct perspective on the toy's location or Sally's incorrect perspective. We predict that performance should improve in the latter condition.

Acknowledgments

We are grateful to the Editor and to three anonymous reviewers for their very informative comments and criticisms. We also gratefully acknowledge the support of grants ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*. The first author further gratefully acknowledges support of a scholarship from Stipendienstiftung Rheinland-Pfalz.

References

1 Wellman, H. *et al.* (2001) Meta-analysis of theory of mind development: the truth about false belief. *Child Dev.* 72, 655–684
2 Onishi, K.H. and Baillargeon, R. (2005) Do 15-month-olds understand false beliefs? *Science* 308, 255–258
3 Southgate, V. *et al.* (2007) Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* 18, 587–592
4 Perner, J. and Ruffman, T. (2005) Infants' insight into the mind: how deep? *Science* 308, 214–216
5 Perner, J. and Roessler, J. (2012) From infants' to children's appreciation of belief. *Trends Cogn. Sci.* 16, 519–525
6 Baillargeon, R. *et al.* (2010) False-belief understanding in infants. *Trends Cogn. Sci.* 14, 110–118
7 Rubio-Fernández, P. and Geurts, B. (2013) How to pass the false-belief task before your fourth birthday. *Psychol. Sci.* 24, 27–33
8 Sperber, D. and Wilson, D. (1986) *Relevance, Communication and Cognition*, Harvard University Press
9 Mascaro, O. and Sperber, D. (2009) The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition* 112, 367–380
10 Buttelman, D. *et al.* (2009) Eighteen-month-old infants show false-belief understanding in an active helping paradigm. *Cognition* 112, 337–342
11 Knudsen, B. and Liszkowski, U. (2012) Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Dev. Sci.* 15, 113–122
12 Southgate, V. *et al.* (2010) Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Dev. Sci.* 13, 907–912
13 Csibra, G. and Gergely, G. (2009) Natural pedagogy. *Trends Cogn. Sci.* 13, 148–153
14 Scott, R.M. *et al.* (2012) False-belief understanding in 2.5-year-olds: evidence from two novel verbal spontaneous-response tasks. *Dev. Sci.* 15, 181–193
15 He, Z. *et al.* (2012) 2.5-year-olds succeed at a verbal anticipatory looking false-belief task. *Br. J. Dev. Psychol.* 30, 14–29