Cognition 130 (2014) 300–308

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

The origins of belief representation: Monkeys fail to automatically represent others' beliefs

Alia Martin *, Laurie R. Santos

Department of Psychology, Yale University, 2 Hillhouse Ave, New Haven, CT 06520, United States

ARTICLE INFO

Article history:
Received 1 April 2013
Revised 19 November 2013
Accepted 20 November 2013

Keywords:
Theory of mind
Belief representation
Comparative cognition
Infant cognition

ABSTRACT

Young infants' successful performance on false belief tasks has led several researchers to argue that there may be a core knowledge system for representing the beliefs of other agents, emerging early in human development and constraining automatic belief processing into adulthood. One way to investigate this purported core belief representation system is to examine whether non-human primates share such a system. Although non-human primates have historically performed poorly on false belief tasks that require executive function capacities, little work has explored how primates perform on more automatic measures of belief processing. To get at this issue, we modified Kovács et al. (2010)'s test of automatic belief representation to examine whether one non-human primate species—the rhesus macaque (*Macaca mulatta*)—is automatically influenced by another agent's beliefs when tracking an object's location. Monkeys saw an event in which a human agent watched an apple move back and forth between two boxes and an outcome in which one box was revealed to be empty. By occluding segments of the apple's movement from either the monkey or the agent, we manipulated both the monkeys' belief (true or false) and agent's belief (true or false) about the final location of the apple. We found that monkeys looked longer at events that violated their own beliefs than at events that were consistent with their beliefs. In contrast to human infants, however, monkeys' expectations were not influenced by another agent's beliefs, suggesting that belief representation may be an aspect of core knowledge unique to humans.

1. Introduction

People understand other agents' behaviors not only in terms of their superficial physical properties, but also as the result of a rich repertoire of unobservable mental states. Much debate in developmental psychology has focused on the problem of how we acquire the ability to represent such unobservable mental states, and how we come to recognize that these mental states can be different from our own. Although early research suggested that children begin to represent others' false beliefs only around four years of age (see reviews in Wellman, Cross, & Watson, 2001), more recent work demonstrates that infants show some understanding of others' false beliefs even in the first two years of life (e.g., Buttelmann, Carpenter, & Tomasello, 2009; Knudsen & Liszkowski, 2012; Kovács, Téglás, & Endress, 2010; Luo, 2011; Onishi & Baillargeon, 2005; Poulin-Dubois & Chow, 2009; Scott & Baillargeon, 2009; Scott, Baillargeon, Song, & Leslie, 2010; Song & Baillargeon, 2008; Song, Onishi, Baillargeon, & Fisher, 2008; Southgate, Chevallier, & Csibra, 2010; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007; Surian & Geraci, 2012). In a landmark paper, Onishi and Baillargeon (2005) observed that 15-month-old infants look longer when an agent with a false belief searches for an object in its true location than when the agent searches in the spot where she saw it last (see also Surian et al., 2007). Infants around this age are also able to take false beliefs into account when inferring an agent's preferences

* Corresponding author. Tel.: +1 (310) 806 2248.
E-mail address: alia.martin@yale.edu (A. Martin).

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.11.016

(Luo, 2011), deciding how to help an agent (Buttelmann et al., 2009; Knudsen & Liszkowski, 2012; Tomasello, 2009), and determining the object of an agent's verbal reference (Southgate et al., 2010).

These new infant findings have been used by many researchers to argue that humans may be equipped with an early emerging system for representing others' beliefs (e.g., Apperly & Butterfill, 2009; Baillargeon, Scott, & He, 2010; Butterfill & Apperly, 2011; Leslie, 2005; Luo & Baillargeon, 2010). Although these specific accounts differ in their claims about the development of the complex belief reasoning observed in human adults, the proposal that infants possess an early emerging system for representing others' beliefs fits with recent "core knowledge" accounts of infant development (see reviews in Kinzler & Spelke, 2007; Spelke, 2004). These core knowledge accounts argue that infants begin life endowed with a set of domain-specific systems for making sense of the physical and social world. These core systems are thought to be older cognitive systems, designed to rapidly solve domain-specific learning problems that our ancestors faced over their evolutionary history. Core knowledge systems are thought to be relatively automatic processes that are constrained by specific signature limits and tend to show characteristic breakdowns under certain situations. These features together require that core knowledge systems show a set of empirical quirks. First, such systems should be experience-independent, and therefore should tend to emerge early in human development. Second, the limits posed by core knowledge systems often persist into adulthood, especially in cases of cognitive load. Finally, because core knowledge systems are thought to be phylogenetically ancient, they are likely to be shared by closely related non-human primates.

To get a sense of these core knowledge features playing out in a different cognitive domain, consider the case of our core knowledge system for object cognition (see review in Kinzler & Spelke, 2007). Some researchers have proposed that there is a core system for representing inanimate physical objects and their movements (e.g., Spelke, Breinlinger, Macomber, & Jacobson, 1992). In line with this view, there is a rich body of evidence that infants possess a set of principles for reasoning about physical objects within the first few months of life, for instance, that objects maintain consistent paths in time and space and tend to cohere (e.g., Aguiar & Baillargeon, 1999; Baillargeon, Spelke, & Wasserman, 1985; Kellman & Spelke, 1983; Kellman, Spelke, & Short, 1986; Leslie & Keeble, 1987; Slater et al., 1990; Spelke, 1990; Spelke, Kestenbaum, Simons, & Wein, 2011; Valenza, Leo, Gava, & Simion, 2006; von Hofsten & Spelke, 1985). Importantly, such principles also implicitly guide adult object processing. When tested on object-based attention tasks, adult participants fall prey to the limits of this system; people fail to track objects that break apart briefly during motion (Scholl & Pylyshyn, 1999) or fail to cohere (vanMarle & Scholl, 2003). Finally, at least some of these principles seem to guide object reasoning in closely related primates (Cacchione & Call, 2010; Flombaum, Kundey, Santos, & Scholl, 2004; Munakata, Santos, Spelke, Hauser, & O'Reilly, 2001; Santos, 2004). These empirical findings together have been used to argue that object knowledge is one of several early emerging core systems for representing the world (see reviews in Santos & Hood, 2009; Spelke, 2004).

Is there empirical reason to argue that a similar core knowledge system exists for early belief reasoning? First, as reviewed above, there is ample evidence that infants begin representing others' beliefs early in life without the need for much experience (e.g., Luo, 2011; Onishi & Baillargeon, 2005; Surian et al., 2007). Indeed, at least one recent study suggests that belief representation may be present in the first few months of life, which is as early as we see evidence for other core abilities. Kovács and colleagues (2010) tested whether 7-month-old infants automatically represent the beliefs of another agent. Their logic was that infants who represent another agent's belief about the location of an object may experience interference in cases in which that agent's belief differs from their own. To test this, Kovács and colleagues presented four different groups of infants with videos involving a cartoon agent who watched a ball roll along a table and behind an opaque occluder. When the ball moved behind the occluder, neither the infant nor the agent could see it. Although each group of infants saw a different series of events, all infants saw a final test outcome in which the occluder was lowered to reveal that there was no ball behind it. Infants in the *Agent and Infant True Belief* condition (ATB-ITB) saw the agent in the video watch the ball roll behind the occluder, out into the open on the table, and then off the visible part of the video screen. In this case, the final test outcome (no ball behind the occluder) was consistent with both the infants' and agent's beliefs. Infants in the *Agent and Infant False Belief* condition (AFB-IFB), in contrast, saw the agent in the video watch the ball roll behind the occluder without rolling back out. In this condition, the final test outcome was inconsistent with both the infants' and agent's beliefs. Indeed, Kovács and colleagues found that infants looked longer in the AFB-IFB condition than in the ATB-ITB condition. In a third condition, *the Agent True Belief Infant False Belief* condition (ATB-IFB), infants saw the agent watch the ball roll off the video screen. Then, while the agent was not watching, infants saw the ball roll back on screen and behind the occluder. In this condition, the final test outcome was inconsistent with the infants' belief even though it was consistent with the agent's belief. Perhaps unsurprisingly, infants looked longer in this ATB-IFB condition than they did in the ATB-ITB condition. In the final and critical *Agent False Belief Infant True Belief* condition (AFB-ITB), infants saw the agent watch the ball roll behind the occluder. Then, while the agent was not watching, infants saw the ball roll out from behind the occluder and off screen. In this condition, the final test outcome was inconsistent with the agent's belief, even though infants should find it expected. Interestingly, infants looked longer in this AFB-ITB condition than in the ATB-ITB case. This pattern of performance suggests that infants reacted to seeing a test outcome that violated the agent's belief even when that outcome was consistent with their own belief. This finding demonstrates that infants (potentially automatically) compute an agent's false belief, even from as young as seven months of age. In this way, infants are able to represent others' false beliefs at around the same time that they