would interesting to know whether such assessments influence their expectations regarding the epistemic reliability of the caregiver and of other people. It is also important to investigate what other strategies children use. For example, at what point are they sensitive to consensus across individuals? Sensitivity to consensus may enable children to differentiate between domains of knowledge that are broadly shared within a community (e.g. language, folk biology) versus domains that are more idiosyncratic or person-specific (e.g. preferences, opinions).

One final point should be emphasized. Children's developing sensitivity to the likely accuracy of one informant versus another might sometimes make them more rather than less vulnerable to misleading information. For example, children who passed false belief tasks made more recognition memory errors if they had been asked misleading questions by an apparently knowledgeable interviewer as opposed to a naïve one [16]. In a similar fashion, children might over-extend their trust in a hitherto reliable informant by endorsing his or her misleading claims.

Philosophers and psychologists alike have commonly treated credulity as one of 'nature's gifts', and therefore likely to be most powerful in childhood. In contrast to the widespread assumption that young children are prone to uncritical credulity, recent evidence shows that they extend trust with appropriate selectivity. Even preschoolers monitor the reliability of a particular informant, differentiate between informants based on the validity of their past claims, and are guided by their interpretation of speakers' minds when evaluating new information from these people. To the extent that theory-of-mind research informs our understanding of mental state attribution, the stage is set for understanding how children make different attributions, depending on what they know, to different individuals, depending on what they claim to know.

References

1 Baldwin, D. and Moses, L. (2001) Links between social understanding and early word learning: Challenges to current accounts. *Social Dev.* 10, 309–329
2 Koenig, M. and Echols, C. (2003) Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition* 87, 179–208
3 Pea, R.D. (1982) Origins of verbal logic: Spontaneous denials by two- and three-year-olds. *J. Child Lang.* 9, 597–626
4 Clément, F. *et al*. (2004) The ontogenesis of trust. *Mind and Lang.* 19, 360–379
5 Koenig, M.A. *et al*. (2004) Trust in testimony: Children's use of true and false statements. *Pschol Sci* 15, 694–698
6 Koenig, M. and Harris, P. Preschoolers mistrust ignorant and inaccurate speakers. *Child Dev.* (in press)
7 Wellman, H. *et al*. (2001) Meta-analysis of theory of mind development: the truth about false belief. *Child Dev.* 72, 655–684
8 Leslie, A. *et al*. (2004) Core mechanisms in 'theory of mind'. *Trends Cogn. Sci.* 8, 528–533
9 Birch, A. and Bloom, P. (2004) Understanding children's and adult's limitations in mental state reasoning. *Trends Cogn. Sci.* 8, 254–260
10 Cheney, D. and Seyfarth, R. (1988) Assessment of meaning and the detection of unreliable signals in vervet monkeys. *Animal Beh.* 36, 477
11 Sabbagh, M. and Baldwin, D. (2001) Learning words from knowledgeable versus ignorant speakers: Link between preschoolers' theory of mind and semantic development. *Child Dev.* 72, 1054–1070
12 Lutz, D. and Keil, F. (2002) Early understanding of the division of cognitive labor. *Child Dev.* 73, 1073–1084
13 Robinson, E. and Whitcombe, E. (2003) Children's suggestibility in relation to their understanding about sources of knowledge. *Child Dev.* 74, 48–62
14 Gopnik, A. and Graf, P. (1988) Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Dev.* 59, 1366–1371
15 Ainsworth, M. *et al*. (1978) *Patterns of Attachment: A Psychological Study of the Strange Situation*, Erlbaum
16 Welch-Ross, M. (1999) Interviewer knowledge and preschoolers' reasoning about knowledge states moderates suggestibility. *Cogn. Dev.* 14, 423–442

1364-6613/$ - see front matter
doi:10.1016/j.tics.2005.08.006

Developmental parallels in understanding minds and bodies

Alan M. Leslie

Department of Psychology and Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854, USA

A recent article by Onishi and Baillargeon presents evidence that 15-month-old infants attribute false beliefs (FBs) to other people. If correct, it lends dramatic new support to the idea that mental state concepts ('theory of mind') emerge from a specialized neurocognitive mechanism that matures during the second year of life. But it also raises new puzzles concerning the FB task – puzzles that have intriguing parallels in results from infants' reasoning about solid bodies.

Tradition has it that horses can't count, children learn to talk by imitation, and infants can't have ideas that are both innate and abstract. Many psychologists just *know* these things. But for twenty years proponents of the Theory of Mind Mechanism (ToMM) have argued that a specialized neurocognitive system, maturing in the first two years in infants, forms the specific innate basis for

Corresponding author: Leslie, A.M. (aleslie@ruccs.rutgers.edu).
Available online 25 August 2005

acquiring a 'theory of mind' [1,2]. The initial motivating observations for this proposal included the following: (i) learnability – the young brain must first notice intangible mental states before it can learn about them; (ii) pretense – the best analysis of early pretend play shows that 18-month-olds possess the concept PRETEND, which is cognate to BELIEF; and (iii) autism – an extensive range of findings support the idea that autistic social development suffers from a specifically impaired ToMM. Still, many psychologists considered the idea of an innate ToMM absurd.

The central dogma in the field quickly came to be that children attribute beliefs to other people only when they pass the 'standard' FB task at 4 years of age [3]. So long as children fail *this* task, they lack the concept of false-belief. Ergo, 'theory of mind' is not innate and there is no such thing as a ToMM. Subsequently, however, a large number of adult brain imaging studies showed what might after all be a neurocognitive system dedicated to 'theory of mind' [4,5]. For a variety of other reasons, faith in the central dogma has been ebbing. Some argue that because of its complexity the 'standard' FB task should simply be discarded as a test of 'theory of mind' [6]. Others have found ways to simplify the task and promote earlier passing [7,8]. Another approach has been to use systematic variations of the task as a tool for developing models of false-belief processing and its limitations [9,10]. Meanwhile, looking-time studies with infants have uncovered ever richer social cognitive skills in the first two years. Starting with Leslie's early work, six-month-old infants were shown to recognize the causal agency of hands [11]. More recently, Woodward showed that infants recognize that active hands pursue goals [12]. Soon after the first year, infants have a surprisingly detailed appreciation of the structure of goal-directed action [13], recognize the epistemic and referential nature of an agent's eye gaze [14], and even what is old and new information for another person [15]. Suddenly, the idea of an early ToMM is not so crazy after all.

The eyes have it

And now comes Onishi and Baillargeon's findings [16]. Fifteen-month-old babies were familiarized with an actor who placed an object in one of two boxes in front of them. The actor then either left the scene (false-belief condition) or remained watching (true-belief condition) while the object moved by itself from the original box into another box (see Figure 1). This recapitulates the essentials of what became the standard 'Sally and Anne' FB task [1] but is entirely non-verbal. Infants' looking times were then recorded to one of two test events: either the actor reached into the original box or into the other box where the object was now. In the false-belief condition, babies looked longer when the actor reached into the current location of the object, shorter when the actor reached into the original box, where she should (wrongly) believe the object to be. In the true-belief condition, infants showed the opposite pattern, looking longer if the actor reached into the original box. Looking times therefore appeared to follow violation-of-an-expectation based on the actor's belief states rather than on object location (Figure 1). These are only the bare bones of the study, which included further conditions controlling for several alternative interpretations.

Other recent findings with infants, mentioned above, mean that Onishi and Baillargeon's results are no longer as isolated as they might have seemed even a few years ago. A torrent of further research will no doubt be stimulated by this result. Defenders of the 'standard' view have already fired their first shot with Perner and Ruffman's commentary on Onishi and Baillargeon [17]. They suggest that infants have 'evolved' either an innate tendency to triple-associate actors, objects and locations or an 'innate rule' that actors look for x where they *last* saw x. It is not clear what advantage 'evolving' such specific tendencies or rules would bring. Certainly neither is evident in three-year-olds who default in the FB task to indicating where the object currently is regardless of actor's history. Nevertheless, ruling out these possibilities will be a spur to further research. Perner and Ruffman also raise so-called 'implicit' FB tasks. These are verbal but incorporate a nonverbal target-of-eye-gaze measure following a verbal prompt, '*I wonder where* [actor] *will look*?' Eye gaze exposes much earlier success, even in two-year-olds [18]. But why, ask Perner and Ruffman, is there no 'implicit' success even earlier? An obvious answer might be the verbal presentation and demands of the 'implicit' task. But it's a good question.

Minds and bodies

It is natural that the conceptual representations produced by an automatic, modular, 'instinctual' mechanism, like ToMM, should be 'implicit', or tacit. Maturation of ToMM in the second year produces early PRETEND- and DESIRE-representations; BELIEF-representations appeared to mature later, in the third year, but Onishi and Baillargeon's landmark findings suggest that this might occur earlier, together with PRETEND and DESIRE. The challenge for future ToMM research is twofold: first, to describe more fully the structure of ToMM; second, to understand why early competence is initially revealed by looking time, then by gaze, and why non-automatic voluntary (e.g. verbal) responses reveal mostly incompetence until much later.

'Theory of mind' research findings now closely parallel those on concepts of physical object solidity. Looking times in infants provide evidence of a very early maturing solidity principle that constrains representations of object mechanics [19]. And yet older infants and toddlers fail to show this knowledge in their manual search behavior and search as if one object can pass through another [20]. Theoretical accounts of FB failures [7,10] and of object-search failures [20] have both appealed to an early competence that is later obscured by failures in the inhibition of prepotent responses. In the case of FB, the prepotent response is to attribute a belief that (in the attributer's eyes) is true. This is a 'best guess' strategy because people's beliefs about everyday matters typically are true. But when it comes to FB, this default response is wrong and obscures competence. Children only succeed on standard FB tasks if they inhibit the default and select the correct false-belief content. Studies of this selection process indicate that it is non-modular, instructable, and the source of the true-belief bias [10]. Intriguingly, Waskett *et al.* (unpublished data) recently showed that two-year-olds in an object-search task fail when searching manually but succeed on eye gaze measures following an

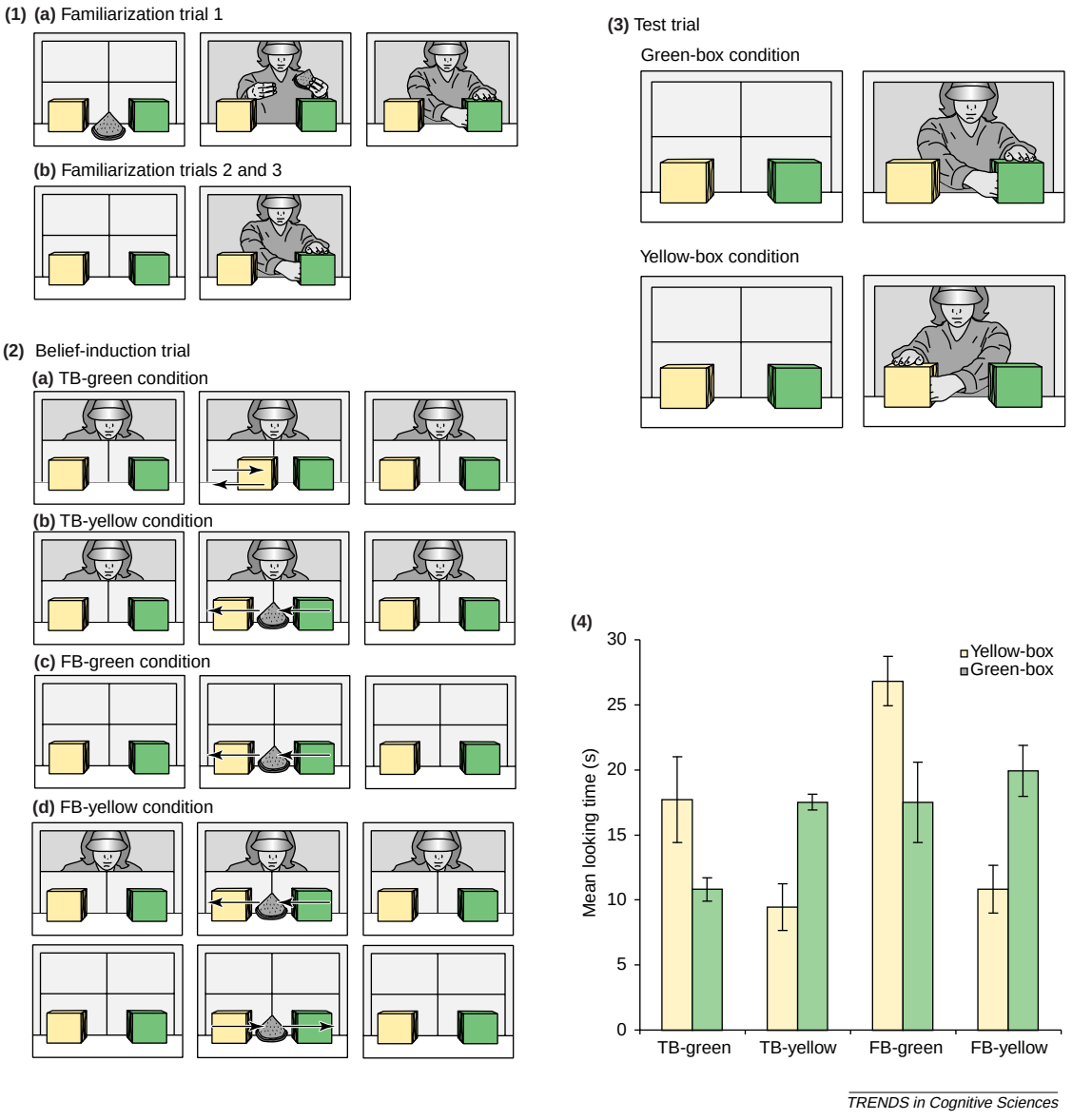

Figure 1. The experimental set-up and results of Onishi and Baillargeon's experiment [16]. **(1)** The initial events watched by all infants. **(a)** shows the actor finding an object and placing it in one of two boxes. In two further trials **(b)**, the actor simply reached inside the green box as though to grasp the object. **(2)** Infants were assigned to one of four conditions determined by the next event they watched. **(2a)** and **(b)** show the two true-belief (TB) conditions in which the actor sees the final position of the object, either remaining in the green box while the yellow box moves (a) or moving from the green to the yellow box (b). **(2c)** and **(d)** show the two false-belief (FB) conditions, in which the actor does not see the final position of the object, either when it moves from the green to the yellow box (c) or when it moves back from the yellow to the green box after the actor had watched it move from green to yellow (d). **(3)** Half of the infants in each condition were then tested on one of two test trials, in which the actor reached into either the green box (Green-box condition) or into the yellow box (Yellow-box condition). **(4)** The mean looking times on the test trial, by condition (error bars show SE). Infants looked longer when the actor reached into a box that the actor should believe was empty, compared with conditions where the actor reached into a box that the actor should believe contained the object. Infants apparently expected the actor's search to be guided by the actor's belief status. (Adapted from [16]).

'*I wonder where*…?' prompt. Thus, toddlers show the same combination of contradictory behaviors in object search as they show a few months later in false belief tasks.

Traditionalists might hope that these patterns and parallels are simply superficial responses to the statistics of surface phenomena, 'evolved' ad hoc 'rules', remarkable coincidences, or a mixture of these. But they are more likely to be signposts to the deeper and tacit roots of conceptual thought. In connecting tacit representations with conscious verbalized reasoning, early brain development might sometimes have to link and reconcile separately functioning systems – as it were, 'repairing' naturally occurring disconnection syndromes.

References

1 Baron-Cohen, S. *et al*. (1985) Does the autistic child have a 'theory of mind'? *Cognition* 21, 37–46
2 Leslie, A.M. (1987) Pretense and representation: The origins of 'theory of mind'. *Psychol. Rev.* 94, 412–426
3 Perner, J. (1991) *Understanding the Representational Mind*, MIT Press
4 Frith, C.D. and Frith, U. (1999) Interacting minds: A biological basis. *Science* 286, 1692–1695
5 Saxe, R. *et al*. (2004) Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.* 55, 04.1–04.38
6 Bloom, P. and German, T.P. (2000) Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77, B25–B31
7 Roth, D. and Leslie, A.M. (1998) Solving belief problems: Toward a task analysis. *Cognition* 66, 1–31

8 Siegal, M. and Beattie, K. (1991) Where to look first for children's knowledge of false beliefs. *Cognition* 38, 1–12
9 Friedman, O. and Leslie, A.M. (2004) Mechanisms of belief-desire reasoning: Inhibition and bias. *Psychol. Sci.* 15, 547–552
10 Leslie, A.M. *et al*. (2005) Belief-desire reasoning as a process of selection. *Cogn. Psychol.* 50, 45–85
11 Leslie, A.M. (1984) Infant perception of a manual pick-up event. *Br. J. Dev. Psychol.* 2, 19–32
12 Woodward, A.L. (1998) Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34
13 Gergely, G. *et al*. (2002) Rational imitation in preverbal infants. *Nature* 415, 755
14 Csibra, G. (2003) Teleological and referential understanding of action in infancy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 447–458
15 Tomasello, M. and Haberl, K. (2003) Understanding attention: 12- and 18-month-olds know what is new for other persons. *Dev. Psychol.* 39, 906–912
16 Onishi, K. and Baillargeon, R. (2005) Do 15-month-old infants understand false beliefs? *Science* 308, 255–258
17 Perner, J. and Ruffman, T. (2005) Infants' insight into the mind: How deep? *Science* 308, 214–216
18 Garnham, W.A. and Ruffman, T. (2001) Doesn't see, doesn't know: Is anticipatory looking really related to understanding of belief? *Dev. Sci.* 4, 94–100
19 Baillargeon, R. (1986) Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month old infants. *Cognition* 23, 21–41
20 Freeman, N. *et al*. (2004) Young children who abandon error behaviourally still have to free themselves mentally: A retrospective test for inhibition in intuitive physics. *Dev. Sci.* 7, 277–282

1364-6613/$ - see front matter
doi:10.1016/j.tics.2005.08.002

Do infants really understand false belief?

Response to Leslie

Ted Ruffman[1] and Josef Perner[2]

[1]Department of Psychology, University of Otago, Box 56, Dunedin, New Zealand
[2]Department of Psychology, University of Salzburg, Hellbrunnerstrasse 34, A-5020 Salzburg, Austria

In his Research Focus ([1] this issue) Leslie argues that the idea of an innate theory of mind module (ToMM) has for too long been considered absurd. We do not consider it absurd but simply think current data can be explained equally well by means of an interaction between genetic endowment and the environment, and that exploration of this interaction should not be curtailed by *a priori* restrictions due to genetic over-specification.

In defense of his position Leslie first cites neuroimaging studies in support of neural structures 'dedicated' to theory of mind. The seeming consensus regarding which brain region is responsible for theory of mind [2] is currently being challenged [3,4] such that different areas are considered responsible for different aspects of theory of mind. Furthermore, it is unclear whether involvement of particular neural regions says anything about innateness. The fusiform gyrus was thought uniquely 'dedicated' to face perception until the discovery of its specialization for cars in car enthusiasts [5]. Is specialized processing of cars also innate? For some functions brain areas become specialized by virtue of experience.

Leslie also cites evidence for theory of mind in infancy. However, even some of the authors cited acknowledge that their results have both a mentalistic and a behavioristic interpretation [6,7]. Indeed, we can distinguish four different explanations:

(1) *Low level similarity to previous encodings*. Leslie does not cogently counter our initial argument [8] against Onishi and Baillargeon [9] along these lines; that cells in the brain code for configurations of persons relating to objects, and infants' looking might indicate these codings.

(2) *Behavior rules* (e.g. 'people look for objects where they last saw them').

(3) *Teleological understanding* [7], in which behavior is understood as being due to goals and external circumstance (true beliefs), and a rationality assumption is made that the most efficient means of achieving the goal are taken.

(4) A *mental understanding* that allows for different perspectives of a circumstance, which is needed for understanding false belief.

Leslie questions why infants would evolve behavior rules. Our points are that such rules are potentially used by primates with which humans share a genetic ancestry [10], that such ancestral abilities also form core theories in other areas that are then modified by means of language [11], or that infants' sophisticated statistical learning abilities [12] would also provide ample means for acquiring such rules. Indeed, the parallels to language acquisition are striking; beliefs that syntax must be innate have been tempered by evidence that infants' statistical learning abilities (which might themselves be innate) permit learning about at least some aspects of syntax [12].

Leslie also argues that 3-year-olds do not use behavior rules in the traditional false-belief task but default to answering in terms of reality. This leaves the question of why their looking to the empty location in implicit tasks [13] does not also default to reality. Instead, such looking is consistent with the use of a rule. It is perfectly plausible that children use this rule in an implicit task but when asked explicitly they use a different strategy. The evidence for 2-year-olds is less clear despite Leslie's claim that they show implicit insight. In the study cited by Leslie [14], only 53% of children (aged 2 yrs 1 mth to 4 yrs 1 mth) looked correctly over the true and false-belief tasks – not compelling evidence for 2-year-olds passing the test. In the

Corresponding author: Ruffman, T. (tedr@psy.otago.ac.nz).
Available online 24 August 2005