Cognition 116 (2010) 297–301

Contents lists available at ScienceDirect

Cognition

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

Brief article

Thinking about false belief: It's not just what children say, but how long it takes them to say it

Cristina M. Atance[a,*], Daniel M. Bernstein[b,c], Andrew N. Meltzoff[c]

[a] *School of Psychology, University of Ottawa, 200 Lees Avenue, Room E228, Ottawa, Ontario, Canada K1N 6N5*
[b] *Kwantlen Polytechnic University, 12666, 72nd Avenue, Surrey, BC, Canada V3W 2MB*
[c] *Institute for Learning and Brain Sciences, University of Washington, CHDD Building, Room 373, Box 357920, Seattle, WA 98195, USA*

ARTICLE INFO

Article history:
Received 29 January 2009
Revised 22 March 2010
Accepted 4 May 2010

Keywords:
Theory of mind
False-belief reasoning
Conceptual development
Response latencies

ABSTRACT

We examined 240 children's (3.5-, 4.5-, and 5.5-year-olds) latency to respond to questions on a battery of false-belief tasks. Response latencies exhibited a significant cross-over interaction as a function of age and response type (correct vs. incorrect). 3.5-year-olds' *in*correct latencies were faster than their correct latencies, whereas the opposite pattern emerged for 4.5- and 5.5-year-olds. Although these results are most consistent with conceptual change theories of false-belief reasoning, no extant theory fully accounts for our data pattern. We argue that response latency data provide new information about underlying cognitive processes in theory of mind reasoning, and can shed light on concept acquisition more broadly.

1. Introduction

False-belief understanding is typically measured using standard tasks such as the Change in Location (e.g., Wimmer & Perner, 1983) or Unexpected Contents (e.g., Gopnik & Astington, 1988). While 3-year-olds often have difficulty with these tasks, most 4- and 5-year-olds do not (Wellman, Cross, & Watson, 2001). For example, when 3-year-olds see that a crayon box contains candles, not crayons, they will state that they originally believed it contained candles, failing to acknowledge their false belief. Verbal responses are informative, but combining them with other measures would enhance our understanding of underlying cognitive processes. One candidate measure – response latency – has a long (Donders, 1969) and rich history in psychological science (Van Zandt, 2000), yet has received little attention in theory of mind research.

At first blush, one might predict that response latencies on false-belief tasks will decrease linearly with age, due to more efficient information-processing capacities. Yet, given existing theories about *why* younger children fail false-belief tasks, one might predict otherwise. We consider predictions from two prominent categories of theories: the conceptual change (e.g., Gopnik & Wellman, 1994; Perner, 1991) and processing accounts (e.g., Fodor, 1992; Leslie & Thaiss, 1992; Roth & Leslie, 1998). According to conceptual change theories, 3-year-olds fail false-belief tasks because they lack an understanding that the mind can misrepresent reality. Children are argued to acquire this understanding around age 4, accounting for their increase in performance on false-belief tasks. In contrast, processing theories situate 3-year-olds' failure on false-belief tasks in their limited processing capacities, not in their lack of an understanding of misrepresentation. With age, processing capacity increases, accounting for older children's higher success rate on false-belief tasks. Both categories of theories lead to unique patterns of predictions about response latencies.

According to conceptual change theories (see Fig. 1a), 3-year-olds appeal to their true belief about the world when responding to a false-belief task, because they do not yet understand that beliefs (theirs or another's) can misrepresent the world and thus what a person believes

---

* Corresponding author.
E-mail address: atance@uottawa.ca (C.M. Atance).

0010-0277/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2010.05.008

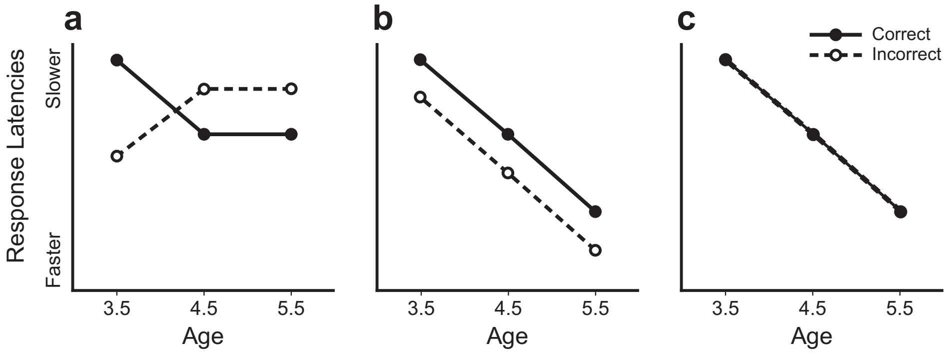

Fig. 1. Correct and incorrect response latencies predicted by conceptual change theories (Panel a), Leslie and colleagues' processing theory (Panel b), and Kikuno et al.'s (2007) processing theory (Panel c). The ages "3.5", "4.5", and "5.5" reflect the mean ages typically used in the literature, and map onto the mean ages in our study.

about the world can conflict with reality (i.e., be "false") (e.g., Wellman, 1990). Accordingly, at age 3, *in*correct responses should be produced quickly – and, importantly, more quickly than correct responses. This is because incorrect responses comport with how 3-year-olds understand the situation – what is in the other's mind must correspond to what is in the world. Conversely, at ages 4 and 5, *correct* responses should take less time to produce than incorrect responses. This is because children now understand mental misrepresentation, and correct responses only require their responding according to this framework. In contrast, incorrect responses would arise if children were unsure and were still actively weighing both the concept they have recently acquired (i.e., "false belief") and the reality of the situation.

Like conceptual change, most processing accounts would predict that incorrect response latencies be faster than correct response latencies at age 3. For example, Leslie, Friedman, and German (2004) argue that an early-developing theory of mind mechanism identifies both "true" and "false" belief contents concurrently. To answer a false-belief question correctly, a "selection processor" inhibits the true belief. This extra processing step should result in longer latencies for correct responses (Carlson & Moses', 2001, executive function account would predict similarly) (see Fig. 1b). This theory appears to make the unique prediction that incorrect response latencies at *all* ages should be faster than correct response latencies, because the former are due to inhibition failures. In contrast, Kikuno, Mitchell, and Ziegler (2007) hold a different processing account. Their account predicts *no* differences in latencies at all ages as a function of correctness (see Fig. 1c), because the reasoning processes underlying correct and incorrect responses are argued to be identical, save for the fact that incorrect responses result from a "reality" bias. Note, however, that by *both* processing accounts, response latencies should become faster with age due to more efficient processing capacities.

Kikuno et al. (2007) tested their account using response latencies with 3- and 4-year-olds.[1] They administered one standard false-belief task and several modified versions in Experiment 1, and only modified versions in Experiments 2 and 3. No differences in response latency were detected as a function of whether children were correct/incorrect on the tasks, and the authors interpreted this null pattern as support for their processing account.

Although Kikuno et al.'s (2007) study provides a good starting point to examine false-belief response latencies, there are several important limitations. First, children only received one standard false-belief task in Experiment 1, yielding one data point for the response latency analyses. This could have contributed to the null effect. Because children's response latencies can be highly variable (Eckert & Eichorn, 1977), multiple data points are preferable. Second, most of the children in their study were 3-year-olds, or young 4-year-olds – an age range that does not capture the progression from systematically failing to passing false-belief tasks. Therefore, their results do not provide a full developmental account of children's false-belief reasoning, nor do they allow us to evaluate the merit of the different theories and predictions that we outlined – our main goal here. Accordingly, the advances we made in this study beyond the previous literature were to assess latencies calculated separately for correct and *in*correct responses in 3.5-, 4.5-, and 5.5-year-olds on a battery of four standard theory-of-mind tasks.

2. Method

2.1. Participants

A total of 240 children participated: 85 3.5-year-olds ($M = 41.75$ months, $SD = 1.72$, Range = 37–47 months; 43 female); 51 4.5-year-olds ($M = 54.35$ months, $SD = 0.60$, Range = 53–56 months; 25 female); and 104 5.5-year-olds ($M = 66.60$ months, $SD = 1.89$, Range = 61–72 months; 47 female). Children completed these tasks as part of three theory-of-mind studies (Bernstein, Atance, Meltzoff, & Loftus, 2007; Sommerville, Bernstein, & Meltzoff, submitted for publication) and so the sample did not include equal numbers of children in each age group. Children came from a large city and were from predominantly middle– and upper–middle-class families.

---

[1] This study was published after we had collected our data. We have, however, accommodated our predictions to include Kikuno et al.'s processing account.

2.2. Materials and procedure

We administered the Unexpected Contents (e.g., Gopnik & Astington, 1988), Change in Location (e.g., Wimmer & Perner, 1983), Occluded Pictures (e.g., Chandler & Helm, 1984), and Appearance–Reality (Flavell, Flavell, & Green, 1983) tasks. These were selected because they test children's understanding that the mind represents reality, rather than being a direct reflection of it, and thus all share an important conceptual core (see Wellman et al., 2001). These tasks were embedded within other tasks not discussed here. For all but the Change in Location task, participants answered two test questions, one about their own false belief/representation and the other about a naïve, same-age peer (a doll) named Ellie. This resulted in a total of seven test questions. Children in the different studies received the tasks and the questions about self and other in either a fixed or counterbalanced order. There were no significant differences in task performance or response latencies as a function of order of task/question administration or study; thus we collapsed the data across these two factors.

In the *Unexpected Contents* task, children were shown a crayon box, asked what they thought was inside and then shown that it contained candles. Children then answered Self ("When you first saw this box, before we opened it, what did you think was inside?"), Other ("What does Ellie think is inside?"), and control ("What is really inside the box?") questions.

In the *Change in Location* task, children saw a character place a ball in a box and then leave. Another character then moved the ball to a cupboard. When the original character returned, children answered experimental ("Now when Billy comes back inside to play, where will he look for the ball?") and control ("Where is the ball really?")[2] questions.

In the *Occluded Pictures* task, children were shown what appeared to be animal ears behind a series of three different windows. Whereas animals were revealed behind the first two windows, for the third window, what looked like bunny ears was really a sunflower. Children answered Self ("When I first showed you this window, all closed up like this, what did you think was underneath?"), Other ("What does Ellie think is under the window?"), and control ("What is really under the window?") questions.

In the *Appearance–Reality* task, children were shown what looked like a rock. After discovering that it was really a sponge, children answered Self ("When you look at this with your eyes right now, what does it look like?"), Other ("What does it look like to Ellie?"), and control ("What is it really and truly?") questions.

For all tasks, children received credit for passing the test questions if they answered the accompanying control question(s) correctly. If control questions were answered incorrectly, the corresponding test question was treated as missing data.

2.3. Scoring

Sessions were video-recorded and response latencies were timed from the videotapes using computer software with accuracy to the hundredth of a second. Children's latencies to respond to the test questions – either verbally or by pointing (in the Change in Location task only) – were timed by a primary coder who scored the videotapes. For a given question (trial), the coder pressed a computer key to initiate timing immediately after the experimenter finished the last word of the test question, and released the key at the onset of the child's response. The computer recorded the duration of the key press for a given trial and these data were stored in a file for subsequent analysis. We excluded latencies if participants answered the control question(s) incorrectly, the experimenter had to repeat the test question, or if latencies were greater than 10 s, which occurred on 7/1300 trials (these values were >3 *SD* from the mean response latencies; inclusion of these data points did not change the overall data pattern). Scoring agreement for latencies was assessed by having a secondary coder time 20% of the data. The two coders had high inter-rater reliability (single measure intraclass correlation = 0.90; 95% CI = 0.89–0.92). We also calculated the mean difference between these two coders by computing the difference between the coders for every trial they scored (*Median* disagreement = 0.10 s; *M* = 0.25 s).

3. Results

Test questions cohered well (*Cronbach's alpha* = 0.78), thus we created a single score by summing individual scores across the questions (*Range* = 0–7). We divided this score by the number of questions for which children passed the required controls, resulting in a proportion score. A one-way analysis of variance (ANOVA) showed that correct responses on the test questions increased significantly with age, $F(2, 237) = 83.31$, $p < 0.001$, *partial* $\eta^2 = 0.41$. Student–Newman–Keuls post hoc comparisons indicated that 5.5-year-olds ($M = 0.78$, $SD = 0.22$) outperformed 4.5-year-olds ($M = 0.61$, $SD = 0.29$), who outperformed 3.5-year-olds ($M = 0.31$, $SD = 0.27$).

Our main interest was differences between correct and incorrect response latencies and the extent to which these conformed to existing theories. Due to the large range in response latencies and the fact that medians are less affected by outliers than are means, we calculated for each participant his/her median latency (collapsed across self and other) for both *correct* and *incorrect* responses. We then entered these median response latencies into a repeated-measures ANOVA with age (3.5, 4.5, 5.5) as a between-subjects factor and response type (correct, incorrect) as a within-subjects factor[3] (note that the ANOVA statistic produces means based on the medians, which are hereafter reported). Results showed that latencies significantly varied by response type, $F(1, 159) = 5.34$,

---

[2] Seventy-four participants received a second control question, "Where did Billy put the ball?".

[3] Participants who answered all seven test questions correctly or incorrectly were omitted from this analysis. This participant exclusion did not affect the overall data pattern. Also, the data pattern remains whether we use median or mean response latency.

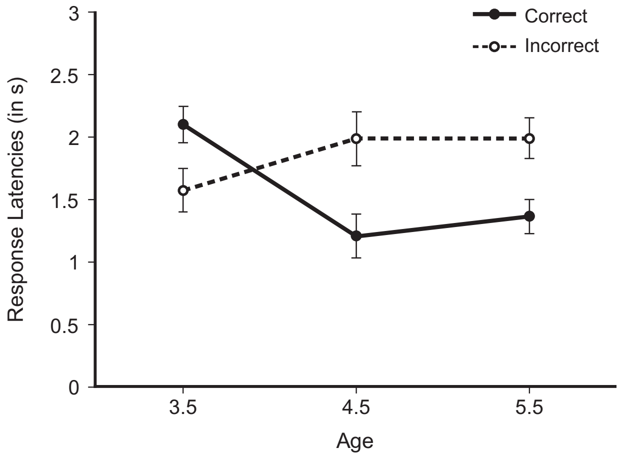

Fig. 2. The effects of age (3.5, 4.5, 5.5) and response type (correct, incorrect) on children's response latencies (in seconds) derived from a repeated-measures ANOVA. Error bars = standard errors.

$p = 0.02$, partial $\eta^2 = 0.03$, but not age $p = 0.43$. The major effect of interest for theory is that there was a highly significant age × response type cross-over interaction for the latencies, $F(2, 159) = 11.33$, $p < 0.001$, partial $\eta^2 = 0.12$ (see Fig. 2). Paired-samples *t*-tests revealed that 3.5-year-olds responded significantly more quickly when *in*correct ($M = 1.59$ s) than when correct ($M = 2.12$), $t(57) = 2.61$, $p = 0.01$. The pattern reversed for 4.5-year-olds (incorrect, $M = 2.00$ s, correct, $M = 1.22$ s), $t(37) = -3.34$, $p < 0.002$, and 5.5-year-olds (incorrect, $M = 2.00$ s, correct, $M = 1.38$ s), $t(65) = -3.10$, $p = 0.003$.[4]

We also ran two between-subjects ANOVAs to better understand the nature of the interaction that we obtained (note that these analyses yield different means from those reported in the repeated-measures ANOVA because they are based on slightly different subjects). Correct response latencies decreased significantly with age ($M = 2.11$, $M = 1.31$, and $M = 1.28$ for 3.5-, 4.5-, and 5.5-year-olds, respectively), $F(2, 212) = 14.17$, $p < 0.01$, whereas incorrect response latencies did not ($M = 1.66$, $M = 1.98$, and $M = 1.97$ for 3.5-, 4.5-, and 5.5-year-olds, respectively), $F(2, 186) = 1.39$, $p > 0.25$. Follow-up *t*-tests, using Bonferroni correction for multiple comparisons, revealed two significant differences: 3.5-year olds were significantly slower than 4.5- and 5.5-year olds when correct ($p < 0.001$ for both tests).

4. Discussion

Our data show a striking cross-over effect: 3.5-year-olds were faster to respond *in*correctly than correctly; the opposite pattern emerged for 4.5- and 5.5-year-olds. The fact that response latencies were shorter for *in*correct than correct responses in 3.5-year-olds is consistent with both conceptual change and processing accounts (with the exception of Kikuno et al.'s, 2007) because both argue that the cognitive processes that are engaged when a 3-year-old passes a false-belief task are different (and more time-consuming) from those when the child fails. However, the fact that older children's correct latencies were *faster* than their incorrect latencies is inconsistent with processing theories. Leslie and colleagues' processing theory (see Fig. 1b) would predict slower latencies for correct responses than incorrect responses across the 3–5 year age range, because correct responses require an extra inhibitory step. Kikuno et al.'s theory (see Fig. 1c) predicts no difference between correct and incorrect latencies across this age range.

By a conceptual change account, however, older children's correct latencies should be faster than their incorrect latencies. This is because correct responses reflect the fluidity with which the older child's mind reasons about false belief – perhaps without the reality-based option being considered prior to responding to a false-belief task (though see Apperly, Back, Samson, & France, 2008; and Cohen & German, 2010, for evidence for and against the idea that false belief reasoning in adults entails processing information about both belief and reality). In contrast, the longer latencies for *in*correct responses suggest that some older children may be experiencing conflict between the correct false-belief-based response and the incorrect true-belief/reality-based option – with the latter ultimately winning out.

It is also important to consider the pattern of correct/incorrect latencies separately across development. Correct latencies decreased significantly between ages 3.5 and 4.5, with no measurable difference between ages 4.5 and 5.5. This finding is inconsistent with processing accounts that posit that children's processing capacity gradually increases with age. Rather, it is consistent with the conceptual change account that argues that, after approximately age 4, children become adept at reasoning about the representational mind. Interestingly, this latency finding does not directly mirror the significant task improvement either between ages 3.5 and 4.5 and 4.5 and 5.5. One possible interpretation is that both 4.5- (61% correct) and 5.5-year-olds (78% correct) have undergone a conceptual shift, indexed by their above-chance performance and, despite accuracy differences, are reasoning about false belief problems in a similar way – and, more importantly, very differently from the 3.5-year-olds (31% correct, below-chance performance).

Incorrect latencies did not significantly change with age, a finding that is inconsistent with both the conceptual change and processing accounts and also the task performance data. Nonetheless, we hypothesize that this stability in latencies may mask different underlying cognitive processes. Whereas 3.5-year-olds' incorrect latencies may reflect consideration of only the reality-based response, we argue that those of the 4.5- and 5.5-year-olds may reflect the consideration of both the reality-based and belief-based options – an interpretation supported by the cross-over interaction that we obtained.

The latency data pertain to debates about conceptual development more generally, which is often described as

---

[4] Although there were differences in response latencies across some of the individual test questions, the fundamental pattern (i.e., 3.5-year-olds answering more quickly when *in*correct than correct, and 4.5- and 5.5-year-olds showing the reverse) emerged across tasks. Specifically, for 17/21 possible comparisons, this pattern held.

a progression from a state of stability to instability, and back to stability (e.g., Carey, 1985; Gopnik & Meltzoff, 1997; Karmiloff-Smith, 1992; Piaget, 1954). Future work might profitably include microgenetic longitudinal studies (e.g., Amsterlaw & Wellman, 2006; Siegler & Crowley, 1992) that trace children's response latencies on theory-of-mind tasks from ages 3 to 5 at frequent intervals (e.g., bimonthly). Response latencies may also prove useful in testing computational models of false-belief understanding (Goodman et al., 2006), which have called for finer-grained quantitative measures of children's understanding. Response latencies are not a magic bullet but provide a much-needed tool that, used in combination with other measures (e.g., children's explanations, neuroscience measures, etc.), will help us chart the processes that underlie the child's developing understanding of mind.

Acknowledgements

We are grateful for financial support from the National Science Foundation Science of Learning Center Grant (LIFE Center, SBE-0354453) and a Standard Research Grant from the Social Sciences and Humanities Research Council of Canada (SSHRC 410-2008-1681). Any opinions, findings, and conclusions expressed in the paper are those of the authors and do not necessarily reflect the views of NSF or SSHRC. Portions of the data were presented at the 2005 annual meeting of the Jean Piaget Society in Vancouver, CA. We thank Allen Thornton and Michael Papsdorf for statistical advice, and Joy Durham, Jill Hyunh, and Bill Peria for help with data coding. David Liu provided helpful comments on earlier drafts.

References

Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognition and Development, 7*, 139–172.

Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition, 106*, 1093–1108.

Bernstein, D. M., Atance, C., Meltzoff, A. N., & Loftus, G. R. (2007). Hindsight bias and developing theories of mind. *Child Development, 78*, 1374–1394.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.

Chandler, M. J., & Helm, D. (1984). Developmental changes in the contributions of shared experience to social role-taking competence. *International Journal of Behavioral Development, 7*, 145–156.

Cohen, A. S., & German, T. C. (2010). A reaction time advantage for calculating beliefs over public representations signals domain specificity for 'theory of mind'. *Cognition, 115*, 417–425.

Donders, F. C. (1969). On the speed of mental processes (W.G. Koster, Trans.). *Acta Psychologica, 30*, 412–431 (Original work published 1868).

Eckert, H. M., & Eichorn, D. H. (1977). Developmental variability in reaction time. *Child Development, 48*, 452–458.

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance–reality distinction. *Cognitive Psychology, 15*, 95–120.

Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition, 44*, 283–296.

Goodman, N. D., Baker, C. L., Bonawitz, E. B., Mansinghka, V. K, Gopnik, A., Wellman, H., Schulz, L., and Tenenbaum, J. B. (2006). Intuitive theories of mind: A rational approach to false belief. In *Proceedings of the twenty-eighth annual conference of the cognitive science society* (pp. 1382–1387).

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

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind* (pp. 257–293). New York, NY: Cambridge University Press.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Kikuno, H., Mitchell, P., & Ziegler, F. (2007). How do young children process beliefs about beliefs?: Evidence from response latency. *Mind and Language, 22*, 297–316.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences, 8*, 528–533.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition, 43*, 225–251.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: The MIT Press.

Piaget, J. (1954). *The construction of reality in the child (M. Cook, Trans.)*. New York, NY: Basic Books.

Roth, D., & Leslie, A. M. (1998). Solving belief problems: Toward a task analysis. *Cognition, 66*, 1–31.

Siegler, R. S., & Crowley, K. (1992). Microgenetic methods revisited. *American Psychologist, 47*, 1241–1243.

Sommerville, J. A., Bernstein, D. M., & Meltzoff, A. N. (submitted for publication). Measuring false belief in centimeters: Adults and children fail to suppress privileged knowledge on a novel change-of-location task.

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin and Review, 7*, 424–465.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

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

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.