UNIVERSITY OF CALIFORNIA

Santa Barbara

Mistaken Identity: Evidence for Demand Characteristics in Tasks Assessing Identity Statements

A Thesis submitted in partial satisfaction of the

requirements for the degree Master of Arts

in Psychology

by

Erin Jamie Horowitz

Committee in charge:

Professor Tamsin German, Chair

Professor Leda Cosmides

Professor Michael Miller

June 2016

The thesis of Erin Jamie Horowitz is approved.

_____
Leda Cosmides

_____
Michael Miller

_____
Tamsin German, Committee Chair

June 2015

Mistaken Identity: Evidence for Demand Characteristics in Tasks Assessing Identity Statements

by

Erin Jamie Horowitz

ABSTRACT

Mistaken Identity: Evidence for Demand Characteristics in Tasks Assessing Identity Statements

by

Erin Jamie Horowitz

Recent research has argued that children's understanding of false belief depends on a conceptual grasp of the distinction between "sense" and "reference". One strand of evidence for this view is a correlation between false belief task performance and tasks assessing the ability to understand a statement indicating that a character described via two different descriptions is in fact the same person ("identity tasks"). The current experiment presented children ranging from 3-5 years of age with a false belief task along with standard versions of this identity task and a modified, "no-conflict" version in which the order of information presented removes a hypothesized pragmatic implication that the different descriptions refer to different entities. Results revealed equivalent performance across all three tasks in younger children, but among older children, the "no-conflict" version of the identity task was performed significantly better than the other two tasks. No reliable association between any tasks was observed in either age group. These results support the conclusions that: 1) identity statement tasks aimed at tapping the "sense-reference distinction" may include pragmatic as well as conceptual demands, and 2) the pragmatic performance demands inherent in identity tasks appear to stem from a different source than performance demands argued to afflict elicited measures of false belief understanding.

TABLE OF CONTENTS

LIST OF FIGURES

I. Introduction

Theory of mind, or the ability to predict and explain others' behavior in terms of their beliefs and desires, is integral to human social interaction. Over the last few decades, considerable debate has surrounded the precise nature of the cognitive mechanisms underlying this ability, particularly in terms of scope. One view, referred to here as the metarepresentational theory, purports that mental states are processed by a mechanism more broadly designed to process metarepresentations (see, e.g., Perner, 1991). An opposing view, referred to here as the theory of mind mechanism (ToMM) theory, advocates an innate mechanism specifically designed to process mental states (see, e.g., Leslie, 1994). The predictions made by each theory have shaped investigations into mental state reasoning across a vast array of areas in psychology.

Within the field of developmental psychology, researchers have largely focused on determining the age at which children are first able to reason about others' false beliefs, thereby demonstrating an understanding that mental states are dissociable from reality (Dennett, 1979). In a typical false belief task (Baron-Cohen, Leslie, & Frith, 1985; Wimmer & Perner, 1983), the child watches as a character, Sally, places a ball in her basket and then leaves the area. While she is gone, her friend Ann removes the ball from the basket and hides it in a box, at which point the child is asked where Sally will look for her ball when she returns. In order to pass the task, children must choose the location where Sally thinks the ball is located, as opposed to the ball's actual location, thus attributing a false belief to Sally that is separate from their own true belief. There exists a well-documented tendency for 3-year-olds to incorrectly select the actual location of the ball when asked to predict Sally's

behavior; this error is typically resolved at around 4 years of age (for review see Wellman, Cross, & Watson, 2001). Not surprisingly, the two competing theories offer different explanations for this behavior, characterized in terms of competence (the conceptual understanding required to pass the task) and performance (the existence of unrelated cognitive skills that are required to pass the task).

Proponents of the ToMM theory argue that conceptual understanding of mental states is masked by other performance demands inherent in the false belief task. Under this model, the ToMM generates specific mental states, and in cases where a prediction or explanation is required, an additional selection processor (the functions of which are not limited to the domain of mental state reasoning) then chooses the most reasonable mental state – typically one that reflects reality (Leslie, Friedman, & German, 2004; Leslie & Polizzi, 2000). Because passing the false belief task requires the selection processor to inhibit its initial choice (the child's own true belief about the location of the ball) in favor of a mental state that does not reflect reality (Sally's false belief about the location of the ball), errors can be attributed to an inability to inhibit an otherwise more rational response, rather than a failure of the ability to process mental states (Friedman & Leslie, 2004; Leslie, German, & Polizzi, 2005). In support of the ToMM theory, past work has demonstrated that decreasing demands on the selection processor by asking children where the protagonist will "look first" increases performance in younger children (Surian & Leslie, 1999), while increasing demands by introducing the protagonist's desire to *avoid* the object decreases performance in older children (Leslie & Polizzi, 2000). Beyond evidence for the masking effects of performance demands, further evidence for an early-developing ToMM has been provided by several studies, which have demonstrated (through measures of looking time and gaze direction) that 13-15 month-old

infants attribute false beliefs to a character with regards to the location of an object (Kovács, Téglás, & Endress, 2010; Onishi and Baillargeon, 2005).

On the other side of the debate, proponents of the metarepresentational theory interpret performance on the false belief task as indicative of a conceptual understanding of metarepresentations (Perner, 1991; Perner, Leekam, & Wimmer, 1987). According to this theory, metarepresentational understanding hinges upon the ability to distinguish *what* is represented from *how* it is represented (Doherty & Perner, 1998). With regards to the false belief task, the child and Sally both have a belief about the location of the ball, but this belief is represented differently in each case: the child believes the ball is in the box, and Sally believes the ball is in the basket. Successful performance thus requires an understanding that the location of the ball (the "what") can be represented in two different senses (the "how"); from there children must intentionally switch to Sally's perspective in order to choose the correct sense (Perner & Roessler, 2012). In a broader sense, the metarepresentational theory predicts a similar performance pattern on false belief tasks and tasks that measure reasoning about non-mental representations, such as photographs and signs (Leekam, Perner, Healey, & Sewell, 2008).

In support of the metarepresentational theory, Perner, Mauer, and Hildenbrand (2011) demonstrated a link between false belief reasoning and identity understanding. In the "lost-and-found" task, a figure in plain clothes is introduced to the child by his profession (e.g., "this is the firefighter") and is immediately hidden from sight. A boy then enters the scene, carrying an item that he thinks someone might have lost. The boy is told the name of the figure who lost the item (e.g., "Mr. Müller"), after which the key *identity statement* – "Mr. Müller is the firefighter" – is provided. The child is then shown two figures – the one that

was introduced at the beginning of the story, and another, unfamiliar figure – and is asked to identify the owner of the item. In line with the sense-reference distinction, a correct answer demonstrated the child's understanding that the two labels ("Mr. Müller" and "the firefighter") refer to the same figure. In testing children ranging from 3 to 5 years of age, Perner et al. found a significant correlation between performance on the lost-and-found task and performance on the false belief task; even when chronological and verbal age were controlled for, younger 3-year-olds were more likely to fail both tasks, and 4- and 5-year-olds were more likely to pass both. The authors interpreted the association between identity understanding and false belief reasoning as evidence for a common underlying metarepresentational capacity, which develops later in childhood (Perner et al., 2011).

Crucially, Perner et al. (2011) have claimed that the lost-and-found task is void of performance demands and thus provides a pure measure of identity understanding. However, note that in the task the child is given two labels for the same figure – one when it is in sight ("the firefighter"), and one when it is not ("Mr. Müller"). Evidence from the literature on language acquisition has shown that young children are biased to assume that labels are mutually exclusive; that is, they assume that objects only have one label, and will consider a newly introduced label as referring to a novel object (Markman & Wachtel, 1988). The lost-and-found task thus appears to be structured in such a way that children are lead to assume that the two labels refer to two different figures. The identity statement ("Mr. Müller is the firefighter") appears against the background of this pragmatic implication, creating a conflict that the child must resolve in order to provide a correct response.

The current investigation proposes that it is the resolution of this conflict, rather than identity understanding itself, which accounts for performance on Perner et al.'s (2011) lost-

and-found task. To test this claim, we implemented a "no-conflict" version of the lost-and-found task, in which the identity statement is provided earlier in the narrative (between the two labels), thereby eliminating the assumption of mutual exclusivity. If the lost-and-found task were free of performance demands, then the removal of this assumption (but preservation of the identity statement) should have no effect on performance. That is, children should exhibit the same performance pattern on the no-conflict version as on the "standard" version of the lost-and-found task. A correlation among all three tasks would support Perner et al.'s argument for a shared reliance on metarepresentational ability. On the other hand, the current investigation hypothesizes that performance will be higher on the no-conflict version, compared to both the standard version of the lost-and-found task and the false belief task. Demonstration of this performance pattern would lend evidence to the existence of a performance demand within the standard lost-and-found task, and would further call into question the shared reliance on metarepresentational ability.

II. Methods

A. Participants

A total of fifty-four children (27 boys, 27 girls) ranging from 3 to 5 years of age were recruited from and tested on-site at three preschools in Santa Barbara, California. Nine children (3 boys, 6 girls) were excluded from the data set: four because they failed to complete the battery of tasks, and five (2 boys, 3 girls) because they missed at least one control question in the false belief task. The final analysis thus included forty-five children (24 boys, 21 girls), who were divided into a younger "3-year old" group (n = 19, mean age = 41 months, SD = 2.7), and an older "4- to 5-year old" group (n = 26, mean age = 55 months, SD = 4.0).

B. Design

Using a repeated measures design, all children received a battery of five tasks: a false belief task, two versions of standard lost-and-found task replicated from Perner et al. (2011), and two versions of the "no conflict" lost-and-found task, which was identical to the replicated task in all aspects save the location of the identity statement within the narrative. The order of task administration was counterbalanced across participants.

C. Materials

All tasks were administered as animated narratives, which were created in Microsoft PowerPoint using Playmobil® figures and presented to participants on an 11.6" laptop. The animations were designed such that experimenters could narrate the story, pause it to ask questions, and continue once the child provided a response. As in Perner et al. (2011), the current study employed four versions of the lost-and-found task, which varied in terms of the name and profession of the man, the building (which corresponded to the man's profession), the name of the boy, the item found, and the animal outside of the building (Table 1). The versions were identical to one another in all other respects. Each participant received all four versions – two for the standard identity condition and two for the no-conflict identity condition. Version order was counterbalanced across conditions.

Version	Animal	Building	Man's Name	Man's Profession	Boy's Name	Item
1	Dog	Fire Station	Mr. Smith	Firefighter	Timmy	Bag
2	Cat	Hospital	Mr. Jones	Doctor	Billy	Hat
3	Bird	Police Station	Mr. Dean	Police Man	Jimmy	Ball
4	Horse	Bakery	Mr. White	Baker	Danny	Keys

Table 1: Labels used in each version of the lost-and-found task.

D. Procedure

Children were tested individually, in an area separated from their main classroom to avoid distraction. At the start of each session, the child was brought into the testing area and seated at a table across from a laptop computer. One experimenter administered the tasks, while a second experimenter sat to the side and recorded the child's responses. For the duration of the session, no feedback was provided as to whether the child provided a correct or incorrect response. Sessions employed all five tasks and lasted between 10 and 15 minutes. After completing the tasks, the child was thanked and walked back to the classroom.

1. False Belief Task

Each child received one false belief task, which was modeled after Wimmer and Perner (1983).

a. Control Questions

Children were asked a number of control questions throughout the task to ensure their comprehension of the narrative. The story opened by introducing two characters – Sara, shown holding a ball and standing next to a green box, and Max, shown standing next to a blue box. After a few seconds, Sara placed her ball in the green box and left the scene. The experimenter then asked the first control question: "Where did Sara put the ball?" *[correct = green box*]. Once Sara had gone, Max removed the ball from the green box, put it into the blue box, and left the scene. The experimenter then asked the control questions: "Where is the ball now?" *[correct = blue box]*, "Who put it there?" *[correct = Max]*, and "Did Sara see that?" *[correct = no]*.

b. Test Questions

After Max left, Sara entered the scene again, wishing to retrieve her ball. The *prediction question* was then asked: "Where will Sara look for the ball first?" *[correct = green box]*. Sara looked in the green box, which was shown to be empty, and the experimenter asked the *explanation question*: "Why did Sara look for the ball in the green box?"

c. Scoring

Children who could not answer all of the control questions correctly were excluded from the final analysis. Children passed the prediction question if they accurately stated where Sara would look for her ball. As in Perner et al. (2011), criteria for passing the explanation question were modeled after Wimmer & Mayringer (1998). Correct responses indicated either the original location of the ball, or that Sara was absent, did not see Max move the ball, or did not know that the ball had been moved. Incorrect responses described the true location of the ball, expressed Sara's desire to find the ball, or took some form of "I don't know."

2. Lost-and-Found Tasks

Children were given two versions each of the standard and no-conflict conditions. Table 2 describes the presentation of the control and test questions in each condition.

a. Standard Condition

This condition was directly replicated from Perner et al.'s (2011) lost-and-found task. At the beginning of the task, a dog was shown standing next to a building (in this example, a fire station). A man in plain clothes entered the scene, was introduced to the child as "the firefighter," and went into the station. A boy then entered, carrying a bag he had found somewhere. He asked the dog whose bag it is, and the dog answered that it belongs to "Mr.

Smith." The dog then provided the *identity statement*: "Mr. Smith is the firefighter." After hearing this, the boy walked up to the station and rang the doorbell. Two men in plain clothes answered: the one who was introduced at the beginning of the narrative, and another, unfamiliar man. The child was then asked the *test question*, "Whose bag is it?" and was required to answer by pointing to one of the figures.

b. No-conflict Condition

The no-conflict condition was exactly matched to the standard identity condition, except for two features. First, to preserve the structure of the identity statement across conditions, here the figure was initially introduced using a proper name ("Mr. Smith") and later described in terms of his profession ("the firefighter").[1] Additionally, and crucially to our study, the identity statement ("Mr. Smith is the firefighter") was presented *between* these two labels, such that by the time the boy in the narrative was told that the bag belonged to the firefighter, the child had already been informed that "Mr. Smith" and "the firefighter" refer to the same figure.

c. Scoring

Children who could not answer all of the control questions correctly were excluded from the final analysis. In both the standard and no-conflict condition, children passed if they correctly pointed to the owner of the item.

[1] In designing the no-conflict condition, preservation of the narrative's structure necessitated either: 1) using the proper name as the initial label (e.g., "This is Mr. Smith" instead of "This is the firefighter"), or 2) reversing the order of labels in the identity statement (e.g., "The firefighter is Mr. Smith" instead of "Mr. Smith is the firefighter"). In either case, reversal of label presentation would have occurred at some point in the narrative; we chose to preserve the identity statement. This reversal could not have imposed an additional performance demand in the no-conflict condition without imposing the same demand on the standard identity condition as originally implemented by Perner et al. (2011).

Step	Condition	
	Standard Identity	No-Conflict Identity
1	A fire station and a dog standing next to it are shown and introduced to the child.	
Control Question	*Where is the dog? [Points to the dog]*	
2	A man in plain clothes enters and is introduced as "the firefighter."	A man in plain clothes enters and is introduced as "Mr. Smith."
Control Question	*Who is this? [the firefighter]*	*Who is this? [Mr. Smith]*
3	The man goes inside the fire station.	
Identity Statement		The child is told, **"Mr. Smith is the firefighter."**
4	A boy enters the scene carrying a bag that he found.	
Control Question	*Where is the bag? [Points to the bag]*	
5	The dog tells the boy that the bag belongs to Mr. Smith. The boy asks the dog where Mr. Smith is and the dog answers, "He's in the fire station."	The dog tells the boy that the bag belongs to the firefighter. The boy asks the dog where the firefighter is and the dog answers, "He's in the fire station."
Identity Statement	The dog adds, **"Mr. Smith is the firefighter."**	
6	The boy walks up to the fire station and rings the doorbell. Two men answer the door.	
Test Question	*Whose bag is it? [Points to the figure introduced at the beginning of the story]*	

Table 2: The sequence of events in the two conditions of the lost-and-found task. Both conditions were modeled after Perner et al. (2011). The identity statement is given after Step 5 in the standard condition, and after Step 3 in the no-conflict condition.

III. Results

For the standard and no-conflict lost-and-found tasks, children were required to provide correct responses to the target question on *both* trials to be considered as passing either condition. For the false belief task, children were required to provide correct responses to *both* the prediction and explanation questions.[2] The impression given by Figure 1 is that a difference in performance across tasks emerges in the 4-year-old group. However, age effects were not significant for either the standard identity condition (χ^2 (1, N = 45) = 1.52, p > .10) or the false belief task (χ^2 (1, N = 45) = 0.53, p > .50). In contrast, performance on the no-conflict identity condition increased significantly as a function of age (χ^2 (1, N = 45) = 15.50, p < .001).

[2] : The analyses in Perner et al. (2011) only included responses to the prediction question. In contrast, the current study combined responses to the prediction and explanation questions to enable comparison of scores on the false belief task to scores on the two lost-and-found tasks; in this way, the child's score for each task was based on two trials. Analysis revealed no significant difference in difficulty between the explanation and prediction questions. Moreover, excluding responses to the explanation question from the analysis did not affect the results reported here.

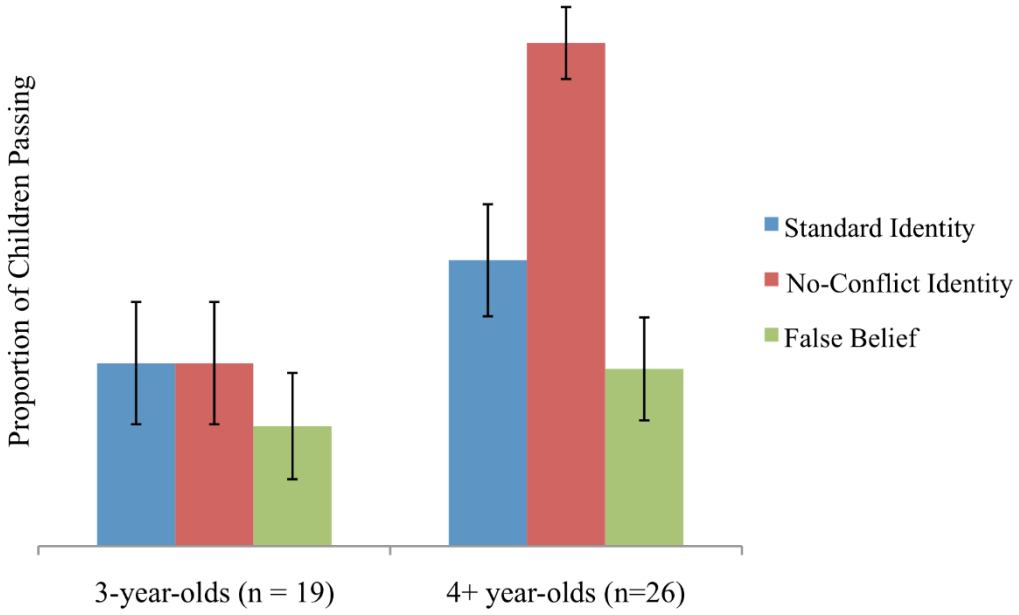

Figure 1: Proportion of children passing each task (Standard Identity, No-Conflict Identity, and False Belief) as a function of age group (3-year-olds and 4+ year-olds). Error bars represent SEM in each age group.

A. Performance on the Two Lost-and-Found Tasks

In examining the younger group, a large proportion of children failed both conditions (42%), while a small proportion passed both (5%). Comparative performance increased in the older group, with no children failing both conditions, and a larger proportion passing both (38%). To assess the relative difficulty of each task in each age group, binomial chi-squared analyses compared the number of children who passed one task and failed the other to the number of children who showed the reverse pattern. Younger children were equally likely (26% in each direction) to pass the no-conflict condition and fail the standard condition, as they were to pass the standard condition and fail the no-conflict condition (McNemar's χ^2 (1, N = 19) = 0.97, p > .50). In contrast, a significant performance difference emerged in the older group: 50% of children passed the no-conflict condition and failed the standard condition, and 12% passed the standard condition and failed the no-conflict condition (McNemar's χ^2 (1, N = 26) = 4.55, p = .021). Taken together, these results suggest that the

no-conflict condition is indeed easier than the standard identity condition, which supports our hypothesis that a performance demand is present in the latter.

To assess the association between the two conditions, a logistic regression analysis was run using performance on the standard identity condition as the dependent variable and performance on the no-conflict condition as the predictor variable. Analysis revealed that performance on the no-conflict identity condition did not significantly predict performance on the standard identity condition ($\chi^2 = 0.61$, $p > .10$), even when age was controlled for ($\chi^2 = 0.23$, $p > .10$)

B. Performance on the Standard Identity Condition and the False Belief Task

The younger group demonstrated generally low performance on the standard identity condition and the false belief task, with a large proportion failing both tasks (53%), and a small proportion passing both (5%). Performance increased in the older group, with fewer children failing both tasks (15%), and a more children passing both (35%). Binomial chi-squared analyses revealed a slight but non-significant tendency for younger children to pass the standard identity condition and fail the false belief task (26%) rather than pass the false belief task and fail the standard identity condition (16%) (McNemar's χ^2 (1, N = 19) = 0.11, $p > .50$). In the older group as well, more children passed the identity condition and failed the false belief task (35%) than the reverse pattern (15%). However, this difference was also non-significant (McNemar's χ^2 (1, N = 26) = .01, $p > .10$).

The apparent closeness in difficulty across age groups suggests that the two tasks share a common underlying ability. To test this possibility, a logistic regression was computed using performance on the false belief task as the dependent variable and performance on the standard identity condition as the predictor variable. In contrast to the results from Perner et

al. (2011), our analysis revealed that the two tasks were not significantly correlated (χ^2 =.002, p = > .50), indicating that performance on the standard identity condition did not significantly predict performance on the false belief task, and lending evidence to the hypothesis that different underlying abilities drive the performance rates on each task. This pattern held even when age was controlled for (χ^2 =.032, p = > .50).

C. Performance on the No-Conflict Identity Condition and the False Belief Task

In examining the younger group, comparative performance between the no-conflict identity condition and the false belief task was consistent with previous analyses; the majority of younger children failed both tasks (53%), while a small proportion passed both (5%). Performance increased in the older group with fewer children failing both tasks (8%), and more children passing both (27%). In terms of relative task difficulty, analyses revealed a slight, non-significant tendency for younger children to pass the no-conflict condition and fail the false belief task (26%) rather than pass the false belief task and fail the no-conflict condition (16%) (McNemar's χ^2 (1, N = 19) = 0.10, p > .50). A significant difference was revealed in the older group: 64% of children passed the no-conflict condition and failed the false belief task, and 4% passed the false belief task and failed the no-conflict condition (McNemar's χ^2 (1, N = 26) = 0.01, p < .001).

Once again, a logistic regression analysis was run to assess the association between the two tasks. Using performance on the false belief task as the dependent variable and performance on the no-conflict condition as the predictor variable, the analysis showed that performance on the no-conflict condition did not significantly predict performance on the false belief task (χ^2 = 0.04, p > .50); this was the case even when age was controlled for (χ^2 = 0.26, p > .50).

IV. Discussion

The aim of the current investigation was to examine the possibility that a performance demand, rather than a lack of metarepresentational ability, influenced performance on a measure of identity understanding. We implemented a standard false belief task, the lost-and-found task as described in Perner et al. (2011), and an additional "no-conflict" version of the lost-and-found task, which was identical to the original version in all respects save the location of the key identity statement within the narrative. The results support our hypothesis that performance demands in the lost-and-found task are responsible for the pattern observed in Perner et al. While task performance was uniformly low in the younger group, differences emerged across the older group, revealing a significantly higher rate of passing the no-conflict condition compared to the standard condition. From this we can infer that the no-conflict condition was much easier than the standard condition. Beyond this, the absence of a significant correlation between the two conditions suggests that they are indexing different abilities.

The relationship between identity and false belief understanding is further called into question by our failure to replicate the correlation found in Perner et al. (2011). Though we found a common level of difficulty between the two tasks, the absence of a significant correlation strongly suggests a difference in underlying ability. Evidence for this is further bolstered by performance on the no-conflict identity condition, which was not correlated with performance on either task. It is entirely likely that the tasks contain different performance demands. While previous studies have implicated the selection processor in false belief tasks (see, e.g., Friedman & Leslie, 2004), the current investigation suggests that failure on the

standard lost-and-found task is caused by the need to resolve a conflict instigated by an assumption of mutual exclusivity.

References

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21, 37–46.

Dennett, D. C. (1979). *Brainstorms*. Hassocks, Sussex: Harvester.

Doherty, M. J., & Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development*, 13, 279–305.

Friedman, O., & Leslie, A. (2004). Mechanisms of belief-desire reasoning: Inhibition and bias. *Psychological science*, 15(8), 547–52.

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

Leekam, S., Perner, J., Healey, L., & Sewell, C. (2008). False signs and the non-specificity of theory of mind: Evidence that preschoolers have general difficulties in understanding representations. *British Journal of Developmental Psychology*, 26, 485.

Leslie, A.M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In *Domain Specificity in Culture and Cognition* (Hirschfield, L. and Gelman, S., eds.), pp. 119–148, Cambridge University Press.

Leslie, A., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind". *Trends in cognitive sciences*, 8(12), 528–33.

Leslie, A., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive psychology*, 50(1), 45–85.

Leslie, A., & Polizzi, P. (2000). Inhibitory processing in the false belief task: Two conjectures. *Developmental science, 1*(2), 247–253.

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

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20(2), 121–157.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255-258.

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

Perner, J., Leekam, S.R., & Wimmer, H. (1987). Three-year-old's difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5, 125-137.

Perner, J., Mauer, M. C., & Hildenbrand, M. (2011). Identity: key to children's understanding of belief. *Science*, 333(6041), 474–7.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in cognitive sciences*, 16(10), 519–25.

Surian, L., & Leslie, A. M. (1999). Competence and performance in false-belief understanding: A comparison of autistic and 3-year-old children. *British Journal of Developmental Psychology*, 17, 141–155.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child development*, 72(3), 655–684.

Wimmer, H. & Mayringer, H. (1998). False belief understanding in young children: explanations do not develop before predictions. *International Journal of Behavior Development*, 22, 403-422.

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