

---

MARCO FENICI

*Università degli Studi di Siena*

*fenici@unisi.it*

---

# WHAT DOES THE FALSE BELIEF TEST TEST?

*abstract*

---

*The age at which children acquire the concept of belief is a subject of debate. Many scholars claim that children master beliefs when they are able to pass the false belief test, around their fourth year of life. However, recent experiments show that children implicitly attribute beliefs even earlier. The dispute does not only concern the empirical issue of discovering children's early cognitive abilities. It also depends on the kind of capacities that we associate to the very concept. I claim that concept possession must be understood in terms of the gradual development of the abilities that underlie the concept in question. I also claim that the last step to possess the concept of belief requires children to understand how beliefs and desires are used in everyday explanations of people's actions. Thus, I suggest that understanding folk psychology as an explanatory theory is what children lack when they fail the false belief test.*

*keywords*

---

*Belief; false belief test; theory of mind; social cognition; folk psychology*

### 1. Introduction: implicit and explicit attribution of beliefs in infancy

The age at which children acquire the concept of belief is a subject of debate. Until recently, psychologists have relied for evidence on a standard procedure, the *false belief test* (FBT). In the classical version of FBT, the *location change* task (Wimmer & Perner 1983; Baron-Cohen, Leslie & Frith 1985), a child is presented with a scene in which a puppet puts an object in a box and leaves the room to play. While the puppet is away, another character moves the object from the box to another place. When the puppet returns, the child is asked: “Where will the puppet look for the [object]?”. If the child points to the box, this demonstrates the ability to consider the puppet’s belief about the object’s old position as opposed to relying on his or her own knowledge of its real position.

Extended research showed that children younger than four perform very poorly on several versions of the test<sup>1</sup>. Not only do they often fail but their scores are rarely above chance (Wellman, Cross & Watson 2001; Wellman & Liu 2004; Milligan, Astington & Dack 2007). We know that this result does not depend on the complexity of the test’s linguistic presentation<sup>2</sup>. Thus, empirical research highlights younger children’s inability to consider others’ point of view. Many have hence argued that children lack the concept of belief before they are able to pass FBT.

Nevertheless, recent data collected through violation of expectancy, gaze-monitoring and other methods challenge this conclusion. Surian, Caldi & Sperber (2007), for instance, found that 13-month infants formed different expectations about a puppet’s future behaviour depending on whether the puppet had previously seen a desired object being put beyond a screen. Therefore, those infants apparently understand the puppet’s knowledge or ignorance about the object’s location. Other studies (Onishi & Baillargeon 2005; Southgate, Senju & Csibra 2007) show that children’s looking time anticipates the right answer in FBT already

<sup>1</sup> Besides the location change task, different version of FBT are the *unexpected content* task (Hogrefe, Wimmer & Perner 1986; Perner, Leekam & Wimmer 1987; Gopnik & Astington 1988), the *hide and retrieve* task (Fodor 1992; Bloom & German 2000), and the *unexpected-identity* task (Gopnik & Astington 1988).

<sup>2</sup> Children’s performance does not improve if they are asked about where the puppet will look for the object, or where it will say, or think, or know that the object is (Wellman *et al.* 2001). Several studies also show that methods to elicit non-linguistic answers from children do not improve the test’s outcome (e.g., de Villiers & Pyers 2002; de Villiers & de Villiers 2000; Wellman, Hollander & Schult 1996).

around the second year of life<sup>3</sup>. Thus, it is argued, children have an *implicit* understanding of belief attribution much earlier than age four. Finally, the same phenomenon has been found using experimental tasks requiring more active choices (Buttelmann, Carpenter & Tomasello 2009). A growing body of empirical evidence hence strongly suggests that children younger than four are already sensitive to others' beliefs although this ability is not manifested in all the contexts in which one would expect it. It appears that, before being able to pass FBT, children have at least an *implicit* understanding of others' beliefs.

In this article, I resolve this apparent inconsistency in the empirical data by contending that the question about the age when children acquire the concept of belief concerns a theoretical, not only an empirical, question: which minimal capacities are necessary to possess a concept? I will claim that children do not generally acquire concepts at once, but that they gradually master them as they acquire various abilities connected to the concepts themselves. Thus, the empirical results that indicate the implicit understanding of others' beliefs may demonstrate only a preliminary, partial possession of the concept of belief, a concept that must be present in a more mature form in order to pass explicit false belief tests. I will argue that the final step in the acquisition of this concept requires children to understand how beliefs and desires are used in everyday explanations of people's actions. Thus, I will suggest, it is the lack of competence in folk psychological explanation that prevents children from passing FBT. I will provide empirical evidence supporting this view.

## 2. Possessing the concept of belief: empirical and theoretical issues

The debate hinges on whether children's implicit belief attribution ability is a case of genuine belief attribution. This is at the same time both an empirical and a theoretical matter. Although the studies just reviewed show that children younger than four already have an implicit understanding of others' beliefs, they do not clarify whether infants already possess the *concept* of belief. Some claim that this is indeed the case. These theorists argue that children's failure on FBT reflects only a performance limitation, which makes it difficult to manifest their competence in reasoning about beliefs in more complex situations (Fodor 1992; Bloom & German 2000). As I will argue below, these theorists are committed to a *minimalist* conception of belief attribution. Others claim that children only possess the concept of belief when they pass FBT. According to them, children's sensitivity to others' beliefs does not demonstrate the possession of the concept of belief, because the relevant competence is so restricted. We are entitled to credit children with the

<sup>3</sup>Clements & Perner (1994) already found that, but the exposition of their methodology was unclear and biased further research (see de Villiers & de Villiers 2003).

concept of belief only when they show the general ability to make explicit predictions about others' behaviour. Call this a *maximalist* conception of belief attribution.

Note that, despite the empirical opposition between the minimalist and the maximalist conception of belief attribution, they share a common theoretical presupposition, viz., that a concept is defined by a set of properties.

Accordingly, concept possession is thought about in terms of all-or-nothing conditions: someone possesses a concept if and only if she is endowed with a particular set of abilities. The debate then concerns whether we are entitled to attribute the concept in question to beings manifesting just a subset of these abilities.

However, the view that concept identity is subject to necessary and sufficient conditions has been widely opposed in both the philosophical (Wittgenstein 1953; Putnam 1975) and the psychological (Rosch 1975) literature. Since the work of Wittgenstein and Quine, philosophers are inclined to think that there are no analytical truths that define a concept<sup>4</sup>. It follows that we neither must expect there is a predefined set of abilities the manifestation of which ensures that one has any particular concept. Instead, the abilities manifesting the possession of a concept are progressively expanded in their domain of possible applications. We may say that each of these improvements reflects a better possession of a concept, but that none of these steps alone marks the acquisition of the ability. Concept possession is a matter of degree.

This conclusion helps to solve the debate about the age when children acquire the concept of belief. The question as it was posed at the beginning of this article is seen to be ill-formed: there is no precise age at which children acquire any concept; *a fortiori*, there is no age when children acquire the concept of belief. However, the question acquires a new meaning once that concept possession is seen along the lines of ability development. Indeed, if concept possession is a matter of degree, it becomes important clarifying (i) which are the relevant abilities for manifesting the possession of the concept of belief, and (ii) how and when they are learned, so that the concept itself is gradually mastered.

Providing a full list of relevant abilities is beyond the scope of this article. Roughly, the list should include, among the others, infants' early abilities to intentionally interpret others' behaviour (Gergely, Nádasdy, Csibra & Bíró 1995), infants' early abilities to participate in joint attention exchanges (Tomasello, Carpenter, Call, Behne & Moll 2005), as well as the early capacities usually associated with implicit false belief attribution. Herein, I want just

---

<sup>4</sup> For a discussion of the problems connected with the traditional theory of concepts see, for example, Fodor (1998, cc. 3-4), and Laurence & Margolis (1999).

to claim that the last ability manifesting the acquisition of the concept of belief will be that of mastering belief attribution not only to predict people's behaviour, but also to explain it. This is a linguistic capacity, and it requires children to master what is usually called common sense, or folk psychology, that is, the theory, which is implicit in common talk about mental concepts, relating reasoning about different mental states.

Now, several abilities are required to master folk psychology as an explanatory practice. On the one hand, children need to understand the syntactic structure of mental verbs, i.e., the fact that mental verbs take sentences (i.e., that-clauses) as their complement<sup>5</sup>. On the other hand, children need to evolve their understanding of the complex network of relations between different mental verbs. Such a task is complicated by the fact that children do not encounter folk psychology all at once, but that they must reconstruct it from their participation in folk psychological narratives (Bruner 1990; Hutto 2008). Limited narrative ability means that children cannot learn to distinguish the functional meaning of different kinds of mental verbs (e.g., thinking as opposed to desiring verbs) in their third year (Rakoczy & Tomasello 2009). If passing FBT were connected to possessing the ability not just to predict, but also to explain others' beliefs, then it would stand to a reason that children cannot pass FBT before they were able to perceive subtle differences between mental state verbs with respect to their role in the explanation of behaviour. In the next section, I will claim that this is indeed the case.

- 3. What does the false belief test test?** In the previous section, I have claimed that possessing the concept of belief is a gradual achievement. The final ability in the relevant suite of abilities constitutive of the possession of this concept, I have claimed, is the ability to use belief attribution in folk psychological explanation. I want now to argue that it is specifically children's lack of a proper explanation of the puppet's behaviour that prevents them from passing FBT before age four. Indeed, in FBT children are explicitly required by verbal interrogation to deliberate about the puppet's behaviour. In many versions of the test, children are asked to justify their prediction, unmotivated results being scored as random answers. But even when no explicit justification is required, deliberating requires one to be able to support her conclusions. Therefore, children try to produce a prediction they can justify in conversational exchange. Here is the impasse, for until age four, children do not master folk psychology as an explanatory theory. Of course, they are able to provide explanations

<sup>5</sup> This ability has been demonstrated to be related to children's performance on FBT (de Villiers & Pyers 2002; de Villiers 2005).

based on the status of the world. Such explanations may be sufficient to explain successful behaviours. However, in FBT, children need to know how beliefs justify behaviour even when behaviour is unsuccessful. Until they lack a general theory about rational agents' behaviour, they are not able to do that. Thus, in the absence of contrasting reasons, what better justification than reality<sup>6</sup>? The only justification they may accept for the puppet's behaviour brings them to a wrong prediction. On the other hand, once they master folk psychology as an explanatory theory, alternative explanations are available; consequently, children even attempt alternative predictions. My proposal is consistent with those accounts according to which children's failure in FBT must be imputed to their lack of linguistic competence (Astington & Baird 2005; Miller 2006; Milligan *et al.* 2007). However, it advances the deeper explanatory hypothesis that it is a lack of explanatory power, rather than some more direct linguistic inability, that explains children's failures on the test. If this is correct, we should find that children's performance in psychological explanations is predictive of their ability to pass the test. In particular, it should not be the case that children can provide reliable psychological explanation without being able to pass it. This is an empirical claim that can be easily tested. An analysis of empirical research in explanatory versions of FBT leads to some *prima facie* conflicting, but nonetheless in the end consistent results. First, let us examine what is apparently contrary evidence. Bartsch & Wellman (1989) found that children's explanatory competence is antecedent to the time when children can pass FBT. Nonetheless, their methodology is questionable with many respects. First, they prompted mental explanations by appropriate questioning when subjects' initial answers did not provide them. Only 39 out of the 79 collected explanations based on false belief attribution were not prompted. However, it is dubious that prompting had no effect on the ratio of mental explanations. Second, they were excessively well-disposed towards children in assessing good responses. Responses like "Because she wants to" including no reference to the object of the desire were scored as correct explanations in terms of belief-desire reasoning. Finally, children's mental explanations were unbalanced towards desires (30% of the unprompted answers) as opposed to beliefs (18% of the unprompted answers)<sup>7</sup>. Batsch and Wellman's result is also odd when compared with much literature

<sup>6</sup> This effect has been somewhere called the "curse of knowledge" (Birch & Bloom 2007). Herein, I am not suggesting that children lack the ability to detach from reality in their predictions. Indeed, they have pretty good predictive capacities, as experiments on implicit belief attribution show. What they lack is an explanation for their own predictive capacities.

<sup>7</sup> A similar analysis is also compatible with Bartsch, Campbell & Troseth (2007), although their method to prompt answers was more ecologically valid.

finding exactly the opposite. For instance, Moses and Flavell (1990) presented three-year-old children with location change scenario. They report that only 3% of children's explanations invoked the character's false belief as the cause of her action; instead, children usually made reference either to the character's desire, or to the outcome of the situation. Similarly, Wellman *et al.* (1996) analysed children's performance in explanatory versions of FBT with the help of thought bubbles. They found that four- but not three-year-olds were able to provide explanations to their answers. Also Wimmer and Mayringer (1998) contrasted children's performance on predictive and explanatory false belief tasks and found that the latter was as just as difficult as the former. They concluded that children do not understand the causal link between misleading informational conditions, epistemic states, and resulting actions. Finally, Atance & O'Neill (2004) reported that only rarely three-and-half-year-olds referred to false beliefs in an explanatory version of an unexpected content task.

Therefore, it seems that children are unable to provide explanations of people's actions until they pass FBT. This supports my claim that FBT does not tap children's general ability to attribute beliefs, but only their more specific capacity to support belief attribution with reliable psychological explanations. Of course, the results provided are not by themselves sufficient to demonstrate such a claim. However, that no negative evidence has been provided already constitutes a positive hint to drive future investigation.

**4. Conclusions** In this paper, I have argued that explaining children's acquisition of the concept of belief requires clarifying our hidden assumptions about concept possession. Now, in the same way as we usually lack definitions to characterise concepts in terms of necessary and sufficient conditions, we also lack criteria to sharply define concept possession, which is better understood in terms of the gradual development of those abilities that possessing a particular concept provides.

If concept acquisition is interpreted as gradual, empirical evidence coming from studies on infants in their second year of life shows that they already possess a minimal concept of belief. However, I have pointed out that the final step to possess the concept of belief requires children to understand how beliefs and desires are used in everyday explanations of people's actions. And I have suggested that it is the lack of this last capacity which prevents children from passing FBT before age four: when children are explicitly asked to predict the puppet's future behaviour, they feel engaged in a conversational practice, and they respond according to the reasons they may provide to support their prediction. But understanding the right reasons for unsuccessful behaviour is a complex task, which is mastered only in

the fourth year. Acquiring it expands children's predictive abilities beyond the limit of contextual interpretation of people's actions. A review of the empirical literature about children's ability to pass explanatory versions of FBT supported this interpretation<sup>8</sup>.

---

<sup>8</sup>I would like to thank Jay Garfield for his insightful comments on an earlier version of this manuscript. I would like to thank the members of the Language Acquisition Group of the Departments of Linguistics at the University of Massachusetts, Amherst, for the discussion about several topics related to this article.



**THE AUTHOR**

Marco Fenici graduated in Logics at the Università degli Studi di Firenze in 2005 with a dissertation on epistemic logics and the problem of logical omniscience. Since 2006, he is a doctoral student in Philosophy and Cognitive Sciences at the Università degli Studi di Siena. His research concerns theoretical and empirical issues about children's acquisition of the concept of belief. During his studies, he has been visiting student at the Technische Universität (Dresden), at the Istituto di Scienze e Tecnologie della Cognizione (Rome), and at the Department of Linguistics of the University of Massachusetts (Amherst).

**REFERENCES**

- Astington, J. W. & Baird, J. A. (Eds.) (2005), *Why Language Matters for Theory of Mind*, Oxford University Press, NY.
- Atance, C. M. & O'Neill, D. K. (2004), *Acting and Planning on the Basis of a False Belief: Its Effects on 3-Year-Old Children's Reasoning About Their Own False Beliefs*. "Developmental Psychology", 40 (6), pp. 953–964.
- Baron-Cohen, S., Leslie, A. M. & Frith, U. (1985), *Does the Autistic Child Have a "Theory of Mind"?*, "Cognition", 21 (1), pp. 37–46.
- Bartsch, K., Campbell, M. D. & Troseth, G. L. (2007), *Why Else Does Jenny Run? Young Children's Extended Psychological Explanations*, "Journal of Cognition and Development", 8 (1), 33.
- Bartsch, K. & Wellman, H. M. (1989), *Young Children's Attribution of Action to Beliefs and Desires*, "Child Development", 60 (4), pp. 946–964.
- Birch, S. A. J. & Bloom, P. (2007), *The Curse of Knowledge in Reasoning about False Beliefs*, "Psychological Science", 18 (5), pp. 382–386.
- Bloom, P. & German, T. P. (2000), *Two Reasons to Abandon the False Belief Task as a Test of Theory of Mind*, "Cognition", 77 (1), pp. 25–31.
- Bruner, J. S. (1990), *Acts of Meaning*, Harvard University Press.
- Buttelmann, D., Carpenter, M. & Tomasello, M. (2009), *Eighteen-Month-Old Infants Show False Belief Understanding in an Active Helping Paradigm*, "Cognition", 112 (2), pp. 337–342.
- Clements, W. A. & Perner, J. (1994), *Implicit Understanding of Belief*. "Cognitive Development", 9 (4), pp. 377–395.
- de Villiers, J. G. (2005), *Can Language Acquisition Give Children a Point of View? In Why Language Matters for Theory of Mind*, "Oxford University Press", pp. 186–219.
- de Villiers, J. G. & Pyers, J. E. (2002), *Complements to Cognition: a Longitudinal Study of the Relationship between Complex Syntax and False-Belief-Understanding*. "Cognitive Development", 17 (1), pp. 1037–1060.
- de Villiers, J. G. & de Villiers, P. A. (2000), *Linguistic Determinism and the Understanding of False Beliefs, in Children's Reasoning and the Mind*, "Psychology Press", pp. 191–228.

- de Villiers, J. G. & de Villiers, P. A. (2003), *Language for Thought: Coming to Understand False Beliefs*, in *Language in Mind: Advances in the Study of Language and Thought*. MIT Press, pp. 335–384.
- Fodor, J. A. (1992), *A Theory of the Child's Theory of Mind*, "Cognition", 44 (3), pp. 283–296.
- Fodor, J. A. (1998), *Concepts: Where Cognitive Science Went Wrong*, Oxford University Press.
- Gergely, G., Nádasdy, Z., Csibra, G. & Bíró, S. (1995), *Taking the Intentional Stance at 12 Months of Age*, "Cognition", 56 (2), pp. 165–193.
- 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 (1), pp. 26–37.
- Hogrefe, G., Wimmer, H. & Perner, J. (1986), *Ignorance Versus False Belief: a Developmental Lag in Attribution of Epistemic States*, "Child Development", pp. 567–582.
- Hutto, D. D. (2008), *Folk Psychological Narratives*. MIT Press.
- Laurence, S. & Margolis, E. (1999), *Concepts and Cognitive Science*, in *Concepts: Core Readings*, "MIT Press", pp. 3–81.
- Miller, C. A. (2006), *Developmental Relationships Between Language and Theory of Mind*, "American Journal of Speech-Language Pathology", 15 (2), pp. 142–154.
- Milligan, K., Astington J. W. & Dack, L. A. (2007), *Language and Theory of Mind: Meta-Analysis of the Relation Between Language Ability and False-belief Understanding*, "Child Development", 78 (2), pp. 622–646.
- Onishi, K. H. & Baillargeon, R. (2005), *Do 15-Month-Old Infants Understand False Beliefs?*, "Science", 308 (5719), pp. 255–258.
- Perner, J., Leekam, S. R. & Wimmer, H. (1987), *Three-Year-Olds' Difficulty With False Belief: the Case for a Conceptual Deficit*, *British Journal of Developmental Psychology*, 5 (2), pp. 125–137.
- Putnam, H. (1975), *The Meaning of 'Meaning'*, in *Mind, Language and Reality*, Cambridge University Press, Cambridge, MA, pp. 215–271.
- Rakoczy, H. & Tomasello, M. (2009), *Done Wrong or Said Wrong? Young Children Understand the Normative Directions of Fit of Different Speech Acts*, "Cognition", 113 (2), pp. 205–212.
- Rosch, E. (1975), *Cognitive Representations of Semantic Categories*, "Journal of experimental psychology: General", 104 (3), pp. 192–233.
- Southgate, V., Senju, A. & Csibra, G. (2007), *Action Anticipation Through Attribution of False Belief by 2-Year-Olds*, "Psychological Science", 18, 587–592.
- Surian, L., Caldi, S. & Sperber, D. (2007), *Attribution of Beliefs by 13-Month-Old Infants*. "Psychological Science", 18 (7), pp. 580–586.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005), *Understanding and Sharing Intentions: The Origins of Cultural Cognition*, "Behavioral and Brain

Sciences”, 28 (5), pp. 675–691.

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

Wellman, H. M., Hollander, M. & Schult, C. A. (1996), *Young Children’s Understanding of Thought Bubbles and of Thoughts*. “Child Development”, 67 (3), pp. 768–788.

Wellman, H. M. & Liu, D. (2004), *Scaling of Theory-of-Mind Tasks*, “Child Development”, 75, pp. 523–541.

Wimmer, H. & Perner, J. (1983), *Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children’s Understanding of Deception*, “Cognition”, 13 (1), pp. 103–128.

Wittgenstein, L. (1953), *Philosophical Investigations*, Blackwell, Oxford.