The classic study by Wimmer and Perner showed that children below the age of about 4 years of age fail to appreciate others’ false beliefs. This opened the floodgates to the theory-of-mind research programme. Recently, a study by Riggs and Simpson has called the central challenge of false belief into question, reporting similar difficulties with true belief. I would argue, however, that not only should the acquisition of false belief remain a central concern but that we should think harder about it and its earlier manifestations.

The scenario has now become a classic [1]. The young child first sees a puppet place a valued object such as a chocolate bar in place A. The puppet then departs. Meanwhile, another story character transfers the chocolate, given a suitable cover story, from place A to place B. The puppet returns hungry. After some control questions, to check that the child has remembered what happened, the child is asked where the puppet will look for his chocolate. Broadly speaking, 4-year-olds say that he’ll look in A, where the puppet put it, and 3-year-olds say he’ll look in B – where it really is. On the face of it, this ‘unexpected transfer task’ demonstrates that the younger children are failing to appreciate not only that agents can have false beliefs but that these cause erroneous behaviour.

The problem is not only with belief, but with non-veridical mental representations more generally. For example, in the ‘appearance–reality task’, 3-year-olds give the wrong answer when asked about ‘What is it really and truly?’ or ‘What does it look like to your eyes?’ when presented with a trick object (e.g. a ‘rock’ that they know to be a sponge really, which they say ‘looks like a sponge’) or an object undergoing an apparent proper change (e.g. a white card placed behind a blue light-filter, which they say ‘is really blue’) [2]. In any event, countless journal articles and not a few books have taken as their premise that non-veridical mental representation in general and false belief in particular do indeed challenge pre-schoolers. We’re all agreed, that is: false beliefs are hard, true beliefs are easy. The question then is why?

The true-belief question
A recent paper by Riggs and Simpson [3] seeks to persuade us that we should mistrust the above premise, because true-belief tasks challenge 3-year-olds about as much as do false-belief tasks. The experiment they did was a variation on the theme of the classic unexpected transfer task. Linda left her book on the floor of the living room before going into the kitchen to fetch a glass of milk. In her absence, her sister took the book upstairs to read. The following true-belief question was asked on Linda’s return to the living room: When Linda left the living room, where did she think her book was?

The basic findings were that (i) 3-year-olds are just as likely to say ‘upstairs’ to the true-belief question as when asked the more usual false-belief question ‘Where does Linda think her book is?’, whereas (ii) they had no difficulty telling the experimenter where the book actually was when Linda left the living room (the ‘memory question’) (see Figure 1).

It is tempting to conclude from this that what makes the true-belief question hard, compared with the memory question, is that it is about belief rather than about reality; and that, accordingly, what challenges young children is simply belief – not false belief in particular. But I think that conclusion should be resisted. The true-belief question does indeed ask about a belief that was true; but the belief is currently false. So what these data do then, rather than overturning our confidence in the specific challenge of false belief, is to encourage us to say the challenge is from ‘currently false’ beliefs. More carefully, the past true belief currently under consideration is currently false; and the child takes its content to be current reality, as in the false-belief task.

![Figure 1. Profile of performance by Riggs and Simpson’s 3-year-olds on false belief (FB) and true belief (TB) tasks. More than 50% of 3-year-old children failed both types of tasks. Redrawn from a table in [3].](image)
In fact, these results are very similar to those from our own laboratory, in a study on children’s reports on what they thought they were doing earlier [4]. We found that 3-year-olds find it just as difficult to report a true belief about what they (or a puppet) were doing (e.g. drawing a handle on a cup) as a false belief (e.g. drawing an ear on a boy). In the former case, the drawing was transformed into one of an ear by overlaying a transparency of a boy over it after it was complete (so the belief that they were drawing a cup handle was true at the time); and in the latter case a cup-drawing transparency overlaid the boy-drawing at the time of drawing so the child (or puppet) falsely believed she was drawing an ear when really she was drawing a cup-handle.

Strong-but-wrong vs. weak-but-right
Both of these experiments bring into sharp relief a central locus of difficulty in the unexpected transfer task and its cognates. Most theory-of-mind tasks set up conflict in children’s minds between:

(a) what they know about the current state of reality (or what is ‘upfront in consciousness’ about it, e.g. in the appearance–reality task, the blue of the card or the sponge nature of the ‘rock’);

(b) a currently non-veridical or not-apparently-veridical representation (e.g. the actual whiteness of the white card held behind the blue filter) in another’s mind or in their own [5].

We might call (a) ‘strong but wrong’ and (b) ‘weak but right’.

What, then, changes developmentally such that older children are able to answer in terms of (b) rather than (a)? There seem to be two main alternatives here:

(i) Children become better at ‘executive inhibition’, meaning, in this context, that they become able to withstand the lure of known reality or of their own salient mental representation [6,7].

(ii) Children acquire a more secure, robust, and adult-like conception of non-veridical and not-apparently-veridical mental representations. This might take the form of a domain-general grasp of the relation between a situation and a representing medium [8], or a domain-specific appreciation of the way in which beliefs are not copies of reality, but representations only ‘in the running’ for truth and to whose truth the agent is committed [9].

It seems to me that we should give both (i) and (ii) a good run for their money. In a sense, the second is more attractive to psychologists because it is inherently more psychological and does not reduce to the view that in this domain children are like highly impulsive adults. Finally, I will give some content to my claim earlier that we should ‘think harder’ about the role of false belief in development.

Earlier origins for false belief?
The ‘hard thought’ I have in mind is that needed to imagine how some, albeit inchoate, appreciation of false belief might be necessary for being a thinker in natural language at all, not only necessary for ‘mind reading’. We need to attend, then, to the very beginnings of language. The philosopher who has given us the requisite food for thought here is Donald Davidson [10]. Davidson believed that to be a thinker in language it is necessary to possess a conception of false belief. This is most clearly illustrated in the labelling of objects. First, a true labeller will have veridicality as her inherent goal: she wants the right word, not just any one that works. Second, the only plausible source of veridicality here is intersubjectivity (true in virtue of social consensus) in the sense that what makes, say, ‘cat’ the true label for these little furry things is that it is the label used by other English-speaking people. Next, how could such a concept of intersubjective truth exist without one of intersubjective error? To illustrate this, Davidson makes use of the notion of ‘triangulation’. This is a primordial joint-attention scenario in which an object is labelled one way by one person (e.g. ‘cow’) and another way by the other person (‘calf’), with mutual monitoring going on between the two people. A true labeller will know what it means to hold a false belief about a label. In fact, the developmental evidence supports Davidson here, in that children as young as 16 months have been shown to respond differently to a human mislabeller (fixating her for longer, and correcting her) than to a mislabelling produced by an audio speaker [11].

Finally, why does appreciation of this kind of false belief develop early whereas that of the standard false-belief task develops later? There are many possibilities, but it might turn out that it is because the first is about social reality (what do people call it?) and the second is about objective reality (where is the chocolate?). In any event, Riggs and Simpson’s paper has the paradoxical effect of reminding us just why the question of how and when children appreciate false belief is centre stage in cognitive developmental psychology.

References