
Alison Gopnik and Virginia Slaughter
University of California at Berkeley, USA
Andrew Meltzoff
University of Washington, Seattle, USA

INTRODUCTION

There is an unfortunate syndrome loose in developmental psychology: Call it “neurotic task fixation.” Those of us interested in children’s developing understanding of the mind have begun to notice the first symptoms of this disease with queasy anxiety. The task in question is, of course, the Wimmer and Perner (1983) false belief task. This chapter and, we hope, this volume is, in part, an attempt to stop the disease before it kills an interesting field yet again. The question should not be “Can three-year-olds do false belief?” Instead, we should ask what aspects of an understanding of belief develop in what ways and in what order, and how these changes are related to one another. How is three-year-olds’ early knowledge of the mind related to their later knowledge? How do precursors lead to the things they precurse?

We have argued that the development of theories of mind involves a genuine conceptual change, or better, a series of genuine conceptual changes. In our formulation, we articulate this as a series of theories each replacing a previous theory (Gopnik & Wellman, 1992). In particular, we and others have proposed a theory change from an early nonrepresentational theory of the mind, with perceptions and desires as the central explanatory concepts, to a later representational theory with propositional contents as central explanatory concepts (Gopnik & Wellman, 1992; Perner, 1991; Wellman, 1990). On the basis of this formulation, we might also make predictions about the origins of an understanding of representation and
about where the earliest precursors to the representational theory are likely to be found.

If this "theory theory" formulation is correct we might expect, indeed we would predict, that the earlier nonrepresentational view of the mind would provide a basis for the later account. The most distinctive feature of a theory account is precisely this systematic relation between earlier and later theories. We would not expect the understanding of false belief, for example, simply to emerge fully fledged when children turn four, with no consistent conceptual relation to earlier accounts, and with no transitional period. In fact, such a developmental pattern would be more consistent with a maturational explanation than with the theory theory. The theory theory predicts that the new theory arises through some combination of relations between the conceptual structures of the old theory and new evidence. What might such structures and such evidence be?

In this chapter we will primarily consider the question of the way that the new representational theory of mind has its roots in the earlier nonrepresentational theory, and also briefly consider the kinds of evidence that might lead to changes in the theory. In both cases we will focus on the ways that an early understanding of perception might serve as a model for the later understanding of belief. We will also present evidence that suggests that children show a better understanding of perceptual misrepresentation than of false belief, and that training children on perceptual tasks can accelerate their understanding of false belief.

There is already evidence that very young children, as young as two-and-a-half years old, have a good understanding of certain aspects of mental life, particularly (1) pretence, (2) desire, and (3) perception. Understanding all three of these types of mental states might be likely precursors to the understanding of belief:

1. Pretence is a likely candidate because, like false belief, it involves acting on a false proposition (Leslie, 1987; see also discussion in Harris, Lillard, & Perner, this volume). Unlike beliefs, however, pretences fail to make reference to the real world. Even two-year-olds pretend and show signs of understanding pretend states (Harris & Kavanagh, 1993) and, in at least one experiment (Flavell, Flavell, & Green, 1987), a pretence-reality task proved to be substantially easier than a parallel appearance-reality task.

2. Understanding desire is also a likely precursor to understanding belief because in adult folk psychology desires, no less than beliefs, are intentional states. Moreover, most commonsense (and philosophical) theories of action see actions as the results of desires and beliefs. The "practical syllogism" (if a person desires x and believes y will lead to x, he will do y) is the first law of folk psychology. Two-and-a-half and three-year-olds show an excellent grasp of the fact that desires may not be fulfilled, that different people may have different desires, and that unfulfilled desires lead to disappointment. They can predict actions by referring to actors' desires and they explain actions in terms of the desires of actors (Aston & Gopnik, 1991; Wellman, 1990; Wellman & Woolley, 1990; Yuill, 1984).

Several studies suggest that placing misrepresentation tasks in the context of earlier desire understanding may also improve performance. In particular, Flavell, Flavell, Green and Moses (1990) found that children were better able to solve a task involving conflicting values, more closely related to desires, than conflicting beliefs. Thus children were better at understanding that Ellie might think that a cookie was yummy whereas they thought it was yucky than at understanding false beliefs. Similarly, Gopnik & Slaughter (1991) found that children were better able to understand changes in their representation of the desirability of objects, than changes in their beliefs. Moses (in press) found that children were better able to understand a misrepresentation task when it was phrased in terms of intentions to act. In all these cases 3-year-olds' performance was not at ceiling, with around 30–40% of children making errors, but it was significantly better than their performance on false belief tasks.

3. A third area of early conceptual understanding, perception, might also serve as an important precursor to understanding belief. Visual perception shares more features with belief than any other mental state. Unlike pretences, visual perceptions make reference to the external world. Unlike desires, they have a mind-to-world direction of fit; the child changes his mind to fit the world, rather than changing the world to fit his mind (see Searle, 1983). Moreover, like beliefs, perceptions are representational states, and so may lead to misrepresentation, in a way that is not true of desire or pretence. You can have an inaccurate or false perception, but you cannot have an inaccurate or false pretence or desire. Much of our adult language about belief inherits perceptual metaphors. We "see the light" or we "change our views." Could children model their understanding of belief on their earlier understanding of visual perception? Could an earlier understanding of perception lead to an understanding of belief by some process of cognitive bootstrapping?

During the recent "theory of mind" boom children's understanding of perception has been less well investigated than their understanding of other mental states. This is a little ironic, since a series of groundbreaking studies on perceptual perspective taking by Flavell and his colleagues were actually among the earliest studies of theory of mind (Flavell, Everett, Croft, & Flavell, 1981; Lempers, Flavell, & Flavell, 1977; Masangkay et al., 1974). For the field as well as for the children, understanding visual perception may have served as a precursor for
understanding belief. There is, however, at least some information about
the understanding of perception that can be gleaned from these earlier
studies and from the current literature.

THE DEVELOPMENT OF AN UNDERSTANDING OF
VISUAL PERCEPTION

Infancy

There is a substantial body of evidence that suggests that a first
understanding of visual perception may have its roots even in infancy.
Even very young infants pay special attention to eyes (Haith, Bergman, &
Moore, 1977; Maurer & Barrera, 1981). A number of writers have pointed
out the significance of joint attention and social referencing behaviours,
which emerge between about nine months and a year, for an understanding
of the mind (Baron-Cohen, 1991; Butterworth, 1991; Wellman, 1993). At
about this age infants develop the ability to follow the point or gaze of
another person towards an object (Butterworth, 1991; Scaife & Bruner,
1975). They begin to use pointing to get someone else to look at an object
(Bruner, 1983; Butterworth, 1991). They also draw attention to objects in
order to discover another person’s attitude towards them. Faced with an
ambiguous situation, such as a visual cliff, infants look to their mother’s
face for information, and may draw her gaze to the objects in question
(Campos & Sternberg, 1980).

These behaviours imply a primitive but genuine understanding of visual
perception. These infants seem to know that the direction of the eyes, or
perhaps just the direction of the head and body, indicates something about
the person’s knowledge of an object. There seems to be some connection in
these children’s minds between the visual display of another person’s head
and eyes, and the phenomenological experience of vision. As a number of
writers have also pointed out, this sort of understanding of vision is also
significant since it implies not only an understanding that people see, but
also an understanding of the relation between people’s vision and the
external world. It implies a primitive understanding of a kind of reference.
These joint attention behaviours are strikingly absent in children with
autism (Baron-Cohen, 1991).

It is worth noting that, although these behaviours are referred to as joint
attention behaviours, they actually imply simply an understanding of joint
perception. In other words, the behaviours clearly suggest that children
know that a particular set of bodily orientations are necessary in order to
see. It is not so clear that they distinguish at this stage between seeing and
attending. This is important, for example, because of Baron-Cohen’s
(Baron-Cohen & Cross, 1992; Baron-Cohen & Ring, this volume) argument
that these behaviours imply some early understanding of propositional
attitudes. If the behaviours really did involve an understanding of attention,
this argument might go through. It is not so clear, however, that simply
attributing vision to another person need involve the attribution of a
propositional attitude.

It is interesting to speculate about the origins of these abilities. Baron-
Cohen and Cross (1992) and Baron-Cohen and Ring (this volume) suggest
that they are the result of a specific innate eye-direction detection module,
and this may be correct. However, they might also have their origins in
particular kinds of experiences in infancy. We have suggested elsewhere
that mutual imitation may be a particularly potent source of information about
the mind (Meltzoff & Gopnik, 1993). In imitation the infant maps an
internal set of sensations, particularly kinaesthetic sensations, onto the
behaviour of another person. For example, the infant maps the felt
sensation of his own tongue onto the visually perceived tongue protrusion
of another person. We have argued that this initial understanding of the fact
that the infant’s felt body is like the seen body of another person might serve
as a framework for the later assumption that the infant’s mind and the mind
of another person are also similar.

Visual perception, like kinaesthetic sensation, is a particularly interesting
intermediate case between understanding bodies and understanding
minds. A wide array of bodily changes will result in consistently different
visual perceptions. Closing your eyes will result in temporary darkness,
moving your head dramatically to one side will cause your visual field to
shift, and so on. The capacity for joint attention seems to presuppose a
mapping from these bodily states (the position of the head and eyes, the
fact that the eyes are open, etc.) onto the phenomenal experience of
perceiving an object.

Neonates have been shown to imitate adult head-turns and may imitate
eye blinks (Meltzoff & Moore, 1989). Normal one-year-olds will imitate
eye closing. In order to do this children must see a correspondence
between their own kinaesthetically registered experience of eye and head
movement and the visual appearance of the eye and head movements of
others. But their own experience, in addition to including kinaesthetic
information about their movements, also includes their experience of the
changes in visual perception that result from those movements. Thus
imitative games might provide a tutorial in the further correspondence
between bodily gesture and experience. I turn my head and see the object
in the corner; I close my eyes and it disappears. Mom performs what I
know is the same gesture, ergo, she must have the same phenomenal
experience.

This account involves a powerful innate mechanism, of course, but it is a
much more general mechanism than a specialised eye-detection module. On
this view the infant innately detects the correspondence of his own own
kinaesthetically perceived body movements and those of others. The infant then generalises this discovery, and assumes that like body movements will be accompanied by like phenomenology, including both kinaesthetic sensations and visual ones. Early imitative abilities, and especially joint imitative activities and games, might help to establish the basic premise of joint attention: When you and I look in the same direction, we see the same thing.

An understanding of visual perception may also be implicated in the development of high-level object-permanence abilities at 15–21 months. Such abilities are commonly seen exclusively as part of the child’s developing understanding of the physical world. But understanding object-permanence also involves knowing the relation between an object’s physical location and its perceptual appearance. The job in an object-permanence problem is to predict where you need to look to find an object when you don’t currently see it. It has as much to do with your perceptual relation to the object as with the object’s physical characteristics. One way of thinking of the highest-level object-permanence abilities is to say that the child discovers that all (or almost all) disappearances are due to changes in the spatial relation between the child’s eyes and the hidden object (see Gopnik, 1984, for further discussion).

The idea that these object-permanence developments reflect a more sophisticated understanding of visual perception is bolstered by data from children’s spontaneous speech. Eighteen-month-olds talk about visual perception in their very earliest language. “Disappearance” words, such as “gone”, are consistently among the very first words to appear (Bloom, 1973; Gopnik, 1982; 1984). These words encode the fact that the child cannot see an object, although she believes that the object continues to exist. They pick out a particular distinctive perceptual relation between the child and the world: Objects are out there but I can’t see them. They do not refer to a particular state of the world itself. These words have consistently been found to be closely and specifically related to the development of high-level object-permanence abilities (Gopnik & Meltzoff, 1984; 1986; Tomasello & Farrar, 1984; 1986). Similarly, two-year-olds use perceptual terms like “see” appropriately and frequently in their spontaneous speech (Bretherton & Beeghly, 1982).

We might see the achievements of infancy as two-fold. First, at around nine months, the child understands that her visual perception may be shared by others (as in joint attention and social referencing). She understands that she and Mom see the same thing. Second, at around 18 months, the child comes to understand something quite general about the reasons why she may fail to perceive an object, even when it exists (as in object-permanence). The child understands that she may not see something from her current point of view.

Eighteen Months to Three Years

These two conceptual achievements of infancy can be combined in complicated ways: I can see something that you don’t, you can see something that I don’t. There is some interesting evidence that suggests that children only begin to understand these more complex relations later in their development. In their very early spontaneous speech, 18-month-old children will only use “gone” when they themselves fail to see an object. They don’t use “gone” to describe the case where they see an object and the other doesn’t, although they may use it when an object is invisible to them, but fully visible to another person (Gopnik, 1984). “Gone” seems to encode a very egocentric notion of disappearance. (It is interesting that it is almost impossible to think of a single adult mental state word that also has this character, that is that only refers to the speaker’s own experience.) By the time children are using other mental states words productively, for example, by the time two- and three-year-olds use “look” and “see,” the words are generalised to refer to the experience of others as well as the child’s own experience.

Lemper et al. (1977) report some fascinating evidence along similar lines. In their studies, very young two-year-olds failed to reproduce situations in which their visual perceptions were discrepant with those of others. When they were asked to show Mommy a picture at the bottom of a long tube, for example, some switched the tube back and forth between themselves and their mother. Similarly, children who were asked to hide an object from their parent behind a screen sometimes produced the egocentric response of hiding the object on the far side of the screen. In studies currently underway in our lab, we have observed similar errors in 18–24-month-olds. One 24-month-old showed an interesting transitional behaviour; she was asked to hide a picture from her mother by turning it so that it faced the child. The child turned the picture back and forth several times, finally settled on the correct perspective (visible to her and invisible to the other) and then walked over to the other side of the table to check that the picture was indeed invisible from that perspective. Initially, very young children may overgeneralise the “social referencing” discovery, and have difficulty in understanding that their vision may not be shared by others.

By two-and-a-half, however, children do seem to understand these aspects of visual perception firmly. That is, they explicitly understand that an object may be visible to them but invisible to another, or vice-versa. Flavell has referred to this phenomenon as “level-1” visual perspective taking (Flavell et al., 1981). Children show a similarly advanced understanding of other perceptual modalities such as hearing and touch (Yaniv & Shatz, 1988). Three-year-old children also are near ceiling in understanding that they may now see an object that they did not see before.
and vice-versa. They understand “level-1” changes in their own visual perspectives (Gopnik & Slaughter, 1991).

Visual-Perspective Taking After Three

The similarity between these “level-1” perceptual abilities and false belief abilities should be clear. In fact, false belief abilities are sometimes referred to as “conceptual perspective taking.” In both cases the child must understand a discrepancy between his own mental state and the mental state of another and, more significantly, must also understand the discrepancy between the mental state and the reality. The object really is there, even though you can’t see it.

However, there are also several different features that distinguish the false belief task from the level-1 perceptual perspective-taking task, which is clearly well-established long before any understanding of false belief. First, the question is phrased explicitly in perceptual terms; it asks what the child sees rather than what the child thinks. Similarly, the task itself is clearly perceptual in character; the change in belief is due to a change in point of view.

In addition, however, the task is a “level-1” task; that is, it requires an understanding that the other, or the past self, simply did or did not know about the object. It requires the child to distinguish between having some knowledge of an object and having no knowledge at all about it. Several theorists have proposed that this simpler understanding requires a different and simpler theory of mind than the understanding of the mind as a representational system; and that this theory is in place at age three. The earlier “level-1” theory has been described in terms of a “copy theory” (Gopnik & Wellman, 1992; Wellman, 1990), “Gibsonian theory” (Aston & Gopnik, 1991), “cognitive connections” (Flavell, 1988), or “situation” view (Perner, 1988; 1991) of belief understanding.

The false belief task, in contrast to the early perceptual tasks, is phrased in terms of “think,” involves a nonperceptual context, and requires that the child understands that the other or the past self represented the object differently, not just that they failed to represent it at all. It requires what Flavell (Flavell et al., 1981) calls a “level-2” understanding and what others have referred to as a “representational model of the mind” (Flavell, 1988; Ferguson & Gopnik, 1988; Perner, 1988). The child who solves these tasks must recognize not only that someone may fail to know something about the world, but also that they may actively misrepresent it. An old saying has it that “it ain’t the things you don’t know that hurt you, it’s the things you do know that ain’t so.” The three-year-olds understand that they or others may not know things; they don’t understand that they may know things that ain’t so.

Which aspect of these tasks is responsible for the difference? Are the perceptual perspective-taking tasks simpler because they are level-1 tasks or because they are perceptual?

A number of tasks seem to fall in between the simple level-1 perceptual task and the false belief task in one way or another. These tasks also seem to be solved at around age three, after level-1 perspective-taking, but before false belief. Wellman and Bartsch (1988) found that three-year-olds, though not two-and-a-half-year-olds, could generally solve what they called a discrepant belief task. In this task, the “level-1” structure remained the same, and the perceptual nature of the task remained the same, but the question was phrased in terms of action (and by implication thought) rather than “seeing.” In this task children were told that objects were in two locations (the drawer and the cupboard) but that a child (say Johnny) had only looked in one location. Children were able to predict that Johnny would only search for the objects where he had seen them. Structurally, this task is quite similar to the “level-1” perspective-taking task; the child must realize that another person only sees part of the world that the child herself sees. Moreover, the task also involves perceptual content. However, the question is phrased in terms of thinking rather than knowing. It is not clear whether the children’s improved performance is due to their “level-1” understanding or their understanding of perception, in particular.

There are some suggestive hints in the literature that even “level-2” perception understanding, that is a genuine understanding of perceptual misrepresentation, might appear before false belief understanding. Flavell kept the perceptual situation and the phrasing in terms of “see,” but changed the task to a “level-2” task; the children were asked about the appearance of an object behind a coloured screen (Flavell, 1978; Flavell et al., 1981; Masangkay et al., 1974). These tasks were more difficult than “level-1” tasks. They require understanding not just that I may see something you don’t, but that the very nature of my perceptual representations may be different from yours. They are the converse of Wellman’s discrepant belief task; they ask about what you see rather than what you think, but they require level-2 rather than level-1 understanding.

These “level-2” perception tasks have often been equated with false belief and appearance-reality tasks. Nevertheless, the data suggest that these tasks may in turn be simpler than these other tasks. In Flavell, Green, and Flavell (1986) there were high correlations between children’s performance on a “level-2” perspective-taking task and on appearance-reality tasks. However, the absolute levels of performance on the visual task were higher than on the A-R task, and higher than typical false belief performance, with over half of the three-year-olds answering correctly. Gopnik and Slaughter (1991) found that children performed substantially better on a level-2 perceptual task asking about their own past perception in a similar “colour change” task,
than on a belief change task. Similarly, Gopnik and Astington (1988) presented children with a false belief task, phrased in terms of thinking rather than seeing, but involving a clearly perceptual task context; in fact, the same “colour change” context in which Flavell had demonstrated level-2 perceptual understanding. Three-year-old children in that study performed best on the task that was phrased in this way, with nearly half the questions answered correctly.

PERCEPTION AS A MODEL FOR BELIEF

All these hints in the literature suggest that children might understand perceptual misrepresentation, either problems phrased in terms of what the child sees, or simply problems that have a plainly perceptual content, some time before they understand false belief. In fact, children might use perceptual misrepresentation as a model for false belief.

How might this take place? We could draw an analogy to other processes of theory change in science, in which the theorist may draw on similarities between the previous theory and the current one. Often the central idea of a new theory is formulated before the fully fledged theory is developed, but is only used in a limited context. Expanding that idea and applying it very generally may be part of the conceptual change that leads to the new theory. We have used the example of Tycho Brahe’s theory of planetary motion to illustrate this (Gopnik & Wellman, 1992). In Brahe’s theory the rest of the planets go around the sun, which in turn goes around the earth. The crucial idea of Kepler’s later theory, heliocentrism, appears, but only in a limited way. Similarly, Darwin’s selection theory depends on ideas that were widely accepted to explain changes in domesticated animals that resulted from breeding. At least part of Galileo’s achievement was to take principles that explained terrestrial movement in a relatively well-understood way and apply them to the movement of the planets.

Similarly, children might take the aspects of their earlier understanding of the mind that are well understood, particularly visual perception, and first formulate an understanding of misrepresentation in that limited context. Later, the idea would be available to be much more widely applied to all cases of belief. In effect, we might imagine the three-year-old psychologist saying “Oh, I get it, these weird belief things are like supercharged perceptions. I know my view of things can change, I guess my views about them can change too.”

UNDERSTANDING PERCEPTUAL MISREPRESENTATION: NEW DATA

In a series of experiments, we have explored children’s understanding of perceptual misrepresentation. In these experiments we have phrased the questions both in terms of perception and in terms of thought, we have posed both “level-1” and “level-2” problems, and we have used a variety of perceptual contexts. We have also asked children about changes both in their own representations and the representations of others.

This last question is extremely important for theoretical reasons. If children are introspecting their own mental states directly, and then using this knowledge as a basis for their generalisations to the states of others, we would expect them to report their own beliefs more accurately than they infer the beliefs of others. On the other hand, if understanding the self and the other develop in parallel, then we might infer that children are using a more general theory of mind to make inferences about both their own mental states and those of others (Gopnik & Wellman, 1992; Gopnik, 1993).

Gopnik and Astington (1988) had found that children’s understanding of their own past states was actually somewhat worse than their understanding of the states of others. Both Wimmer and Hartl (1991) and Moore, Pure, and Furrow (1990) found no difference. Gopnik and Slaughter (1991) also found advanced levels of performance on desire and perception tasks for the self task that were similar to the levels reported by Flavell on “other” tasks. Again, however, there has been no systematic comparison of these conditions for the same children. A systematic pattern in which easy tasks were easy both for self and other, and difficult tasks were difficult both for self and other, would provide particularly convincing evidence that the two types of knowledge develop in parallel.

**Experiment 1**

In the first experiment we tested 14 children from preschools in the Berkeley area. Children’s ages ranged from 3:2 to 3:11, with a mean of 3:7. The children were given two perception tasks, both involving a table with two chairs on either side, one blue and one brown, and a free-standing screen in the middle. In each case the child entered the room and immediately sat on the blue chair, on one side of the screen. The experimenter asked the child to identify what was on the table, and then asked the child to move to the other chair and again identify what was on the table. The experimenter then called in a confederate adult, Mamie, who sat in the original blue chair facing the child, and then the experimenter asked the child the test questions.

In the level-1 task the screen on the table was opaque and two different objects were on either side of it, a spoon on one side and a cup on the other. Thus children saw only one object and then the other, and Mamie similarly saw only one object. Children were asked “What do you see?” when they sat on the blue chair and “Now what do you see?” when they moved to the brown chair. For the “see” question the children were asked “What did you/ Mamie see on the table when you/she were over there, on the blue chair? Did
you see a cup or did you see a spoon?" This is equivalent to the "level-1" perceptual perspective-taking task. For the "think" question children were asked "What did you/Mamie think was on the table, when you/she were over there on the blue chair? Did you think there was just a spoon, or also a cup?" This question is analogous to the Wellman and Bartsch "discrepant belief" question.

In the level-2 task the screen was cut out of coloured red translucent plastic. A free-standing green paper cut-out of a cat was placed on the far side of the screen. The result was a cat that looked black when it was first viewed and then turned out to be green when the speaker’s visual perspective was changed. This condition was similar to Flavell’s "level-2" perspective-taking task. As in the previous experiment, children were asked "How does the cat look to you? Does it look green or does it look black?" and then "Now how does the cat look to you...?" For the "see" question they were asked "How did the cat look to you/Mamie when you were over there, on the blue chair? Did it look like a green cat or did it look like a black cat?" For the think question they were asked "What did you/Mamie think was on the table when you/she were over there, on the blue chair? A green cat or a black cat?"

Children were also given a standard "Smarties and pencils" "deceptive box" false belief task. Children were shown a closed candy box that was then opened to reveal pencils. Children were asked "What do you think is inside this box?" both before and after they opened the box, to ensure that their belief had indeed changed. Mamie entered the room, and saw the closed box. Children were then asked "When you/Mamie first saw the box, all closed up like this, what did you/she think was inside it? Did you/she think there were candies inside or did you/she think there were pencils inside?"

At the beginning of each testing session, children were given two control tasks to ensure that they understood the questions, and could remember the past events. In the perception control task, children sat in the blue chair and saw one object on the table, then they moved to the other side of the table and the experimenter removed the original object and replaced it with another object. As in the perception task, children were asked to report what was on the table before and after they moved chairs. Then children were asked "What was on the table when you were over there on the blue chair?" In the belief control task, children were shown a box with one content, the content was removed, and another content was placed in the box. Again children were asked to report what was in the box, both before and after the change. The children were then asked "When you first saw the box, all closed up like this, what was inside it?" Children were only included in the study if they answered both these questions correctly.

Table 8.1 shows the results. There was no overall difference between performance on the self and other questions, with 67% of the self questions correctly answered, and 65% of the other questions correctly answered. There were also no significant differences on each individual task. We therefore combined the two answers to give children a single score: 0, 1 or 2.

Children performed significantly better on the perception tasks, including both “see” and “think” tasks and level-1 and level-2 tasks, than on the belief task (for level-1 see vs. belief, t = 8.64, P < 0.001; for level-1 think vs. belief, t = 5.70, P < 0.001; for level-2 see vs. belief, t = 4.43, P < 0.001; for level-2 think vs. belief, t = 6.62, P < 0.001). Moreover, the absolute level of performance on the perception tasks, with around 70–80% of the questions answered correctly, is also considerably higher than the usual false belief performance. There were no significant differences among the perception tasks. However, as we might expect from the earlier literature, the level-1 “see” task did show the highest performance, with 89% correct, near ceiling.

The results of this first experiment did indeed appear to support a difference between perceptual and nonperceptual tasks. In particular they replicated, in a single experiment, the pattern of results garnered from the previous literature. As in Wellman and Bartsch, the level-1 think task for the other proved simpler than the false belief task. Notably this was also true for the question about the child’s own beliefs, which had not previously been tested. As in Flavell et al. (1981) and Gopnik & Slaughter (1991), level-1 perceptual perspective taking was near ceiling, and easier than false belief. The novel and surprising finding, however, was that the level-2 perceptual misrepresentation tasks, both using “looks like” and “think,” were also easier than the false belief tasks.

However, some points of difficulty arose. First, the fact that the children received the “see” questions as well as the “think” questions about the same materials may itself have improved their performance on the “think” questions. In other words, asking children to say explicitly what they saw may have influenced their reports of what they thought. This would itself suggest, of course, that children may be using perception as a model for belief.
Second, there were slight differences in the form and procedure for the belief and perception questions. In two previous experiments (Lewis & Osborne, 1990; Siegal & Beattie, 1991), using a particular form of question improved false belief performance. In other studies, however, changes in form, including the use of the particular form in the Lewis and Osborne study, had no effect on performance (Gopnik & Astington, 1988; Hala, 1991; Moore et al., 1990; Robinson & Mitchell, 1992). In Experiment 1, the control tasks ensured that all the children understood the question form. Nevertheless, this possibility seemed worth controlling for.

Finally, in a number of cases, apparently improvements in performance in one study have failed to emerge in further experiments (see earlier; also Hala, Chandler, & Fritz, 1991 vs. Sodian, Taylor, Harris, & Perner, 1991; Robinson & Mitchell, 1992 vs. Robinson & Mitchell, in press). This suggests that sampling variation may sometimes lead to misleading differences in performance on variations of the task. It seemed particularly important, therefore, to replicate any such finding with an additional sample of children and different materials and questions.

**Experiment 2**

We tested an additional 12 3-year-olds from a local daycare centre in Berkeley. Ages ranged from 3:1-4:3, with a mean of 3:8. Children received the level-1 and level-2 perception tasks and the belief task exactly as in Experiment 1. They also received the two control tasks. In this experiment, however, children were not asked explicitly to state what they saw or thought before and after the change, as in Experiment 1. This helped ensure that the children did not discuss what they saw at the time that they answered the “think” question. Moreover, they were not asked “see” and “think” questions about the same materials, and they were not asked the level-1 “see” question, the easiest question, at all. Instead, they were only asked “think” questions about the table and the cat, and they received an entirely different level-2 “see” task, more closely analogous to the “belief” task. In this task, children were shown a stick immersed in water so that it appeared to be bent. The stick was removed from the water so that children could see it was straight. The stick was then returned to the glass, and an adult, Craig, entered the room. The children were asked “When Craig/you first saw this stick, in this glass here, how did the stick look to him/you? Did it look like a straight stick or did it look like a broken stick?” Thus the question was phrased in the same way as the belief question, and involved a similar sequence of events. In other respects the procedures were identical to the previous experiment.

Results are shown in Table 8.2. As in the previous experiments, there were no significant differences between performance on the “self” and “other” questions on any of the tasks. The general levels of performance on the perception tasks were somewhat lower in this experiment, with 66% of the level-1 questions correct and 79% of the level-2 questions correct. This suggests that the explicit use of the “see” question in Experiment 1 may have had a facilitatory effect, though the difference did not reach significance. Even without an explicit “see” question, however, the levels of performance on the perception tasks were still higher than those in previous studies using the false belief task. Moreover, as in the previous experiment, the children’s actual performance on the perception tasks was consistently better than on the belief task (for level-1 think, t = 7.33, P < 0.001; for level-2 think, t = 9.94, P < 0.001). The new level-2 “see” task with the stick, which used exactly the same phrasing as the belief task and involved a similar sequence of events, also proved to be easier than the belief task. Correct answers were given to 62% of the questions in this task, and performance was significantly higher than in the belief task (t = 4.91, P < 0.0005).

The previous experiment suggested that children performed better on perception tasks than belief tasks, even when they were not asked “see” questions about them, and even when they were not asked the easiest level-1 “see” question. It was still possible, however, that the fact that children also received a separate level-2 “see” task influenced their performance on the level-2 “think” task. Moreover, we wanted to investigate yet another question form, and to ensure that the question form was identical in the perception and belief tasks. We had already done this in Experiment 2 for the level-2 “see” task; in this experiment we constructed a similar question for the level-2 “think” task.

**Experiment 3**

We tested 18 children recruited from local preschools. Ages ranged from 3:1 to 4:3, with a mean of 3:7. The “other” questions were asked about a small stuffed Snoopy puppet, which was placed on the correct chair. Only “think” questions were asked, ensuring that the children were not drawing the
analogy to the “see” questions explicitly. As in Experiment 1, children were asked to state explicitly what they thought before and after the change. They were asked “What do you think is in the box?”, “What do you think is on the table?” and “What colour do you think the cat is?” The belief question was “When you/Snoopy first saw the box, what did you/he think was inside it then? Did you/he think there were candies inside or pencils inside?” The level-1 perception question was “When you/Snoopy first saw the table, what did you/he think was on it then? Did you/he think there was just a spoon on the table or also an apple?” The level-2 question was “When you/Snoopy first saw the cat, what colour did you think it was then? Did you think it was green or black?” A similar form was used for the control questions.

Table 8.3 shows the results. These children performed better on the false belief task, with 47% of the questions correctly answered. Performance on the level-1 and level-2 “think” tasks was almost identical to the performance in the previous study, with 63% and 77% correct, in spite of the wording change. As in the previous study, the crucial level-2 think question was significantly easier than the belief question (t = 2.01; P < 0.05). The level-1 think question (analogous to the Wellman discrepant belief task) also showed improved performance compared to false belief, although this difference failed to reach significance. Importantly, this was true even though the three questions were virtually identical. Finally, as in the previous experiments, there was no significant difference between level-1 and level-2 versions of the “think” question.

The larger N and improved false-belief performance in this study also allowed us to look at the younger and older children separately. Some earlier studies found that the false belief task was most difficult for the younger three-year-olds. This was also the case in the current study. The 9 children under 3:6 only answered 33% of the false belief questions correctly, whereas they answered 66% of the level-1 perception questions and 72% of the level-2 perception questions correctly. The contrast between perception and belief was most striking for these younger children (t = 3.95, P < 0.01 for level-1 vs. belief; t = 4.72, P < 0.01 for level-2 vs. belief).

### Table 8.3

<table>
<thead>
<tr>
<th></th>
<th>Level-1 Think</th>
<th>Level-2 Think</th>
<th>Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>12</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>% Correct</td>
<td>63</td>
<td>77</td>
<td>47</td>
</tr>
</tbody>
</table>

As in the previous study, there were no significant differences between performance on the “self” and “other” questions. Moreover, the larger N in this experiment allowed us to test whether there was a positive association between answers on the “self” and “other” questions. Did children who answered one of these questions correctly also answer the other correctly? We can use the statistic phi to evaluate the association of two dichotomous variables. The phi values for the relation between self and other performance were 0.71 (P < 0.01 by Fisher’s exact test), 0.68 (P < 0.05 by Fisher’s exact test), and 0.41 (P = 0.10) for performance on the belief, level-2 think and level-1 think questions respectively.

This set of experiments suggested that children were better able to understand perceptual misrepresentation than false belief. Children were particularly competent when the perceptual nature of the task was made explicit (by saying “see” or “looks like”) but they also showed improved understanding even when the tasks were phrased in terms of “think,” in other words when the test question was identical to the classic false belief question. The actual levels of performance were strikingly similar to the levels on desire tasks in earlier experiments, with about 60–75% correct. This suggests that children may initially understand misrepresentation in the context of desire and perception and only later extend that understanding to belief.

## Inducing Conceptual Change: A Training Study

These experiments, however, only address half of the story. We have hypothesised both that the earlier theory may provide a basis for the new theory, and that new evidence causes the earlier theory to be restructured (Gopnik & Wellman, 1992). If understanding perceptual misrepresentation really contributes to the later understanding of false belief, then letting children experience the similarities between these two areas of understanding might help to accelerate conceptual change. Giving children information or experience with perceptual misrepresentation should improve their performance on false belief tasks.

Moreover, more generally on the theory theory, conceptual change should be a consequence of the accumulation of relevant counter-evidence to the theory. We have suggested that there might be three stages in theory change (Gopnik & Wellman, 1992). In the first stage, children might simply ignore countervailing evidence. There is substantial evidence to support this claim in the literature on false belief. In particular, Moses and Flavell (1990) found that explicitly presenting children with negative feedback, in the course of the false belief experiment itself, had no effect on their answers. Similarly, children who explicitly express their earlier belief in the
representational change task may say that they actually said, as well as thought, that there were pencils in the box (Wimmer & Hart, 1991).

In a second phase, children might take into account counter-evidence but only when they were forced to by the experimenter. Bartsch and Wellman (1989) found that some children who were forced to explain counter-evidence to the false belief task invoked false belief though they did not use the concept predictively. Similarly, Mitchell and Lacocke (1991) found that reminding children of a past explicit action they had performed (choosing a picture of Smarties), which contradicted the theory, improved performance on a representational change task.

Presumably, at some point these explanations are not only applied in the specific context of the falsification but are much more generally applied to all cases of false belief. This last phase was what we were interested in the last experiment. It is not too surprising that children can be induced to give the right answer in the task itself by a combination of feedback and cues; actually it is more surprising that children resist such feedback. The question is whether such counter-evidence might have a longer-lasting effect on their understanding of the task. Therefore we provided feedback in the tasks but only tested its effects days or weeks after the training was accomplished, and on a different task. Moreover, if the theory theory is correct, training the children on related tasks—tasks that do not explicitly concern false belief but call on the representational theory in similar ways—should also accelerate false belief understanding.

We have gathered such data in a training study. Three-year-old children were pretrained on a standard “deceptive box” false belief task (candles in a crayon box), both for self and the other. Only children who passed a control task and yet failed the belief task were included in the study. Three groups of children, with ten children in each group, were randomly assigned to one of three training conditions. In each condition the children were visited twice in the course of the two weeks following the pretraining visit. In the control condition children were given a conservation of number task on each visit. The tasks were designed to be very similar in overall length and structure to the false belief tasks. In each case children saw an array of objects and were asked, “Which line has more?” The array was rearranged and children were asked “Now which line has more?” Children received feedback on their responses, being told “Yes, this line had more” or “No, the other line had more.”

In the belief training condition children received different false belief tasks with different materials on each visit. In one task children saw golf balls, which turned out really to be soap. In the second task children turned the pages of a book. On each page of the book a small cut-out hole in the black overlay revealed the ears of an animal; the whole animal was visible when the overlay was turned. On the last page, the object underneath the overlay turned out to be a flower, and the “ears” turned out to be petals (see Gopnik & Astington, 1988). In each case children were asked “What did you/Snoopy think these were? Ears/golf balls or petals/soap?” They received positive and negative feedback about their responses. The experimenter corrected the child’s “self” response, “No, you thought they were golf balls” and the Snoopy puppet corrected the “other” response, “No, I thought they were golf balls.”

In the third condition, the most interesting one from our current perspective, children in the intervening sessions received a perception task and a desire task, both designed to tap an understanding of misrepresentation. The perception task was the level-2 “cat” task, phrased in terms of “see.” In the desire task children were seen at snack time and were asked “Do you want an apple slice?” Then they were fed apple slices until they were no longer hungry. They were asked “Did you want an apple slice?” and told “Snoopy hasn’t had a snack, does he want an apple slice?” (see Gopnik & Slaughter, 1991). Just as in the belief tasks, the children received differential feedback about their responses from Snoopy or the experimenter.

Finally, in a post-test session exactly two weeks after the original test, children received yet another false belief task. This time the task involved a band-aid box with a small book inside. The task was administered in exactly the same way for all three groups of children and in exactly the same way as in the original pretest.

Table 8.4 shows the results. Children in the belief and perception/desire conditions performed significantly better on the post-test than those in the control condition, both for self and other. This is particularly striking, since the perception and desire tasks did not explicitly mention the word “think” or raise the false belief problem at all.

**CONCLUSION: DEVELOPING THEORIES OF MIND REALLY DO DEVELOP**

The results we report here suggest an (at least) four-stage process in children’s developing theory of mind. (In fact, other data suggest even earlier developments in infancy and toddlerhood.)

![Table 8.4](image)

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Belief</th>
<th>Perception/Desire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>3</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>% Correct</td>
<td>13.6</td>
<td>54</td>
<td>40.9</td>
</tr>
</tbody>
</table>

174  GOPNIK, SLAUGHTER, MELTZOFF

8. CHANGING YOUR VIEWS 175
1. In the earliest phase, before two-and-a-half, children begin to develop a foundational, if sometimes egocentric, nonrepresentational understanding of perception, as revealed in their joint attention, social referencing, and object-permanence abilities.

2. At two-and-a-half or so, children show a firm understanding of nonrepresentational states such as perception and desire. These children can solve simple desire tasks (Aasting & Gopnik, 1991; Wellman & Woolley, 1990) and level-1 perspective-taking tasks. They cannot yet understand more complex tasks such as discrepant belief tasks (Wellman & Bartsch, 1988) or level-2 perception tasks (Flavell et al., 1981).

3. At around age three, children begin to show more competence on these tasks. They begin to show a better understanding of representational aspects of desire, as in the Flavell et al. (1990) values task and our desire change task (Gopnik & Slaughter, 1991). They also begin to show an understanding of level-1 perception, even when the question is phrased in terms of belief, as in the discrepant belief task, and of level-2 perspective-taking, that is perceptual misrepresentation, as in the tasks described here. Children’s performance on all these tasks is not at ceiling, but more than half of the time they answer these questions correctly. Moreover, at least some of the time, they can invoke preliminary accounts of misrepresentation if they are confronted by direct counter-evidence (Bartsch & Wellman, 1989; Mitchell & Lachée, 1991). All of these abilities are interesting conceptual precursors to the fully fledged representational theory of belief, which emerges at around age four. Pieces of the later theory—the very idea of belief, and the idea of perceptual misrepresentation and change and diversity of values—start to appear.

4. Finally, we suggest, at around age four, the force of the accumulating counter-evidence to the earlier theory leads the child to generalise the notion of misrepresentation from perceptual contexts, and to develop a general, predictive, and widely applicable notion of false belief.

This account makes sense of a variety of pieces of evidence about early false-belief-like understanding. It seems to us to do so considerably more coherently than alternative “information-processing” accounts. These accounts propose that the central representational model of the mind is in place at age three or earlier but that various more superficial information-processing and task difficulties mask that competence (Fodor, 1992; Lewis & Osborne, 1990; Siegal & Beattie, 1991). Unfortunately, the issue has become focused on the single question of whether or not children can “do better” on false belief tasks; the neurotic task-fixation we referred to earlier. Such improved performance, by itself, tells us little of interest. In particular, it does not tell us that an information-processing account is correct. The question is, why are there these differences in the

performance? Do they reflect the performance demands of the tasks or their conceptual content?

In fact, to confirm one or the other account we would need to look at a wider pattern of evidence. If the information-processing account is correct, we would want evidence that a range of tasks with similar information-processing demands, but widely different conceptual content, were similarly difficult for three-year-olds and easy for four-year-olds. For example, three-year-old children might always have difficulty answering questions phrased in a particular way. Similarly, we would want evidence that systematically varying information-processing complexity, in predictable ways, had effects on the developmental pattern. In contrast, on the theory theory account, tasks with very different information-processing demands, involving different materials, questions, etc., but with similar conceptual content, would be expected to develop in concert. Variations in conceptual content, with similar or even identical task demands, might lead to different patterns of development.

In terms of the three- to four-year-old shift, the pattern of results is more congruent with the second picture than the first. The results of the current study, as well as many other results in the literature, suggest that three-year-old children have little difficulty dealing with questions and tasks identical to the false belief task when the conceptual content of those tasks doesn’t require an understanding of representation. This is dramatically true for the myriad control tasks that are standardly used in testing false belief. Indeed, the very logic of these tasks is to control for information-processing factors.

It is also true for the intermediate perception and desire tasks. Children in our studies, as well as in Wellman and Bartsch (1988) and Flavell et al. (1990), could answer questions literally identical to the false belief question, in tasks of indistinguishable performance complexity, when they concerned perception, level-1 understanding, or desire. These findings by themselves suggest that it would be difficult to identify a general information-processing difficulty, present at age three but not at age four, that could explain these developments. Why would these difficulties not apply to the control tasks, or to the intermediate tasks? In fact, no coherent account along these lines has ever been proposed, nor have there been any studies that showed that non-theory-of-mind tasks, with similar task demands, developed in close concert with false belief.

In contrast, there are now a number of studies suggesting that tasks with similar conceptual content, but very different information-processing demands—particularly tasks involving a representational understanding of belief—do develop in concert with false belief. These include not only the false belief and representational change tasks, but other tasks such as appearance-reality tasks (Flavell et al., 1986), tasks measuring children’s understanding of the sources of their beliefs (Gopnik & Graf, 1988; O’Neill
& Gopnik, 1991; O'Neill, Astington, & Flavell, 1992), and subjective probability tasks (Moore et al., 1990), as well as analyses of the child's spontaneous speech (Wellman, 1990). All these measures suggest significant changes between ages three and four. Moreover, in several studies there have been significant and high correlations between performance on many of these measures, even with age controlled (Flavell et al., 1986; Gopnik & Astington, 1988; Moore et al., 1990).

The complementary alternative side of this pattern is also beginning to emerge. Evidence that information-processing changes by themselves cause differences in performance is thin on the ground and these results are typically fragile, weak, and variable from study to study. In contrast, changing conceptual content can have robust effects on tasks that are otherwise very similar to the false belief task. The Wellman & Bartsch (1988) discrepant belief task, confirmed here, is a striking example of this. So is the children's ceiling performance on pretence, imagery, and level-1 perception tasks, in contrast to their intermediate performance on desire, intention, and perception tasks that call on representational understanding, and their poor performance on tasks that involve a representational understanding of belief (Gopnik & Slaughter, 1991). Similarly, forcing children to confront counter-evidence to their particular theory seems to induce changes in understanding, as seen in Bartsch and Wellman (1989), Mitchell and Lacochee (1991), and in our current training study. This is a result more congruent with the theory theory than with information-processing theories.

The results we report here, along with other results in the literature, suggest a truly developmental account of developing theories of mind. Understanding is not simply something that emerges fully blown from the child's head, like Athena from the head of Zeus, and then, also rather like Athena, wraps itself in the fogs and disguises of performance constraints. Instead understanding is a process, something painfully earned and constructed. Children move from confident ignorance at two-and-a-half to the glimmers, false starts, and puzzles of the three-year-old and finally to the fully productive and explanatory model at age four. Lest this sound too triumphant, the new model will itself lead to new puzzles, contradictions, and false starts, such as the puzzle of understanding inference or second-order belief (Wimmer, Hogrefe, & Solian, 1988; Perner, 1991). The children trying to understand our minds are, in this regard, very much in the same fix as the psychologists trying to understand theirs.

REFERENCES


