Chapter Seven

Applications of the Account

In chapter five I declared my intention to develop an account of belief that has practical utility for working philosophers and psychologists. To have practical utility, an account must promote clear thinking on the topic at hand, it must help its users make sense of current research, and it must direct their attention away from fruitless inquiries into more productive ones. I believe that the account presented in the previous chapter has this kind of practical utility. The reader has seen the utility of the account in handling the many examples of "in-between" believing presented in that chapter; but to see the real value of the account for philosophical and psychological research, it is necessary to see how the account interfaces with actual contemporary research in these fields.

In this chapter, I will apply my dispositional account of belief to four areas of current research, two in philosophy and two in developmental psychology. We will see philosophers and psychologists repeatedly stumble over the kinds of in-between cases of belief that have been the focus of my attention in these chapters. And we will see energy directed away from useful avenues of inquiry into counterproductive attempts to squeeze genuinely mixed cases of believing into simple all-or-nothing descriptions.

1. Two Philosophical Puzzles

I will begin by describing two philosophical puzzles into which I think we can gain insight by application of my account. I will then show how my account applies to these puzzles and other potentially troublesome similar cases.

Kripke's Puzzle about Belief

The first of the two philosophical puzzles I will be discussing here is put forward in Saul Kripke's (1979) paper, "A Puzzle about Belief". In this paper, Kripke describes several cases in which he thinks standard assumptions about belief lead to paradox. The most fully fleshed-out of these problem cases is that of Pierre, a native French speaker who does not know English, but who grows up reading travel guides and hearing tales of the beauty and magnificence of a certain distant town called 'Londres'. If someone were to ask Pierre, in French, whether he thought that town was pretty, he would assent, and it seems quite natural to say that he believes that London is pretty. Later in his life, Pierre moves to London without knowing it is the same town he calls 'Londres', and he thinks it an ugly place. He would heartily assent to the English sentence, 'London is not pretty'. At the same time, since he has not learned that 'Londres' is the French word for 'London', he would still be willing to claim, in French, that 'Londres est jolie'. He thinks, in other words, that 'Londres' and 'London' name different places, the first pretty and the second not.

Now, Kripke argues, we are on the edge of paradox. If we take Pierre's French utterances seriously, we seem to be compelled to say that Pierre thinks that London is pretty, 'London is pretty' being the English translation of the French sentence to which Pierre sincerely assents. On the other hand, if we take Pierre's English utterances seriously, we seem compelled to say that he thinks that London is not pretty. So Pierre would appear to have contradictory beliefs. Even, however, if we are comfortable describing people as having contradictory beliefs in some cases, in Pierre's case the matter is especially strange: He would seem to be guilty of no logical error but simply a lack of information. It seems unfair to convict him of logical inconsistency.

Can we escape the difficulty by denying either (a.) that Pierre believes that London is not pretty or (b.) that Pierre believes that London is pretty? Rejecting the first claim seems pretty much out of the question: Pierre lives in London and sincerely says that it is not pretty. Rejecting the second claim is a little more tempting. Perhaps Pierre no longer believes that London is pretty. Certainly he did once believe this. He and his French buddies dreamed of someday visiting the beautiful town they called 'Londres' and read about in travel books. But if he did once believe that London is pretty, then ought we not allow that he still believes it? He will still assent to all the same claims, expressed in French, to which he would have assented as a youth. If he ran into his old French buddies, they would see in his eyes not disgust but the familiar dreamy glaze as he

talked about someday visiting the beautiful town of 'Londres'. If everything he ever learned in England were rubbed from his brain, the memories and opinions he still has from France would be amply sufficient to ascribe him the belief that London is pretty.

Note that if we take Pierre's English utterances seriously and say Pierre does not believe that London is pretty and then turn around and say that Pierre does believe that London is pretty on the basis of his French utterances, it is not only Pierre who has contradictory beliefs, but we ourselves.

So what does Pierre really believe about London? Does he really believe it is pretty, or does he really believe it isn't? Or can we make sense of the claim that Pierre really believes both? Or does he, perhaps, have *no* beliefs about London's beauty? In the face of apparently decisive objections to all these options, Kripke announces that the puzzle here is a genuine puzzle, on a par with such famous philosophical puzzles as the Liar's Paradox.

A small body of literature has grown up in response to Kripke's puzzle. Richard Garrett (1991), elaborating on an earlier suggestion by Hilary Putnam (1979), argues that Kripke's puzzle shows that all our beliefs about any object must be qualified by identifying knowledge that allows us to uniquely single out that object. We should not say that Pierre has any bare, unqualified beliefs simply *about London*. Rather, Pierre believes that London, identified in whatever way he associates

with the name 'Londres', is pretty; and he believes that London, identified in whatever way he associates with the name 'London', is not pretty. So long as his associations with the name 'Londres' are not the same as his associations with the name 'London', Pierre's beliefs are not contradictory, and the puzzle disappears.

Appealing as this solution might seem, it has difficulties. First, we should note that it is one thing to say that in order to believe anything about London we must have some identifying knowledge of it; it is quite another thing to claim, as Garrett does, that this knowledge is implicit in and qualifies all our other beliefs about the city. You and I may both believe a lot of things about London that don't uniquely identify it -- such as that it is a big city in England with red double-decker buses and good Indian food -- but if I identify London as the largest city in England and you identify it as the capitol of England, none of my beliefs can, on Garrett's view, be either the same as or inconsistent with any of yours. If I claim that London is pretty and you claim that it is not pretty, we have not, despite appearances, contradicted each other. Since each statement is qualified by different identifying knowledge, neither statement, by Garrett's own assertion, entails the denial of the other. Surely this is a rather counterintuitive position to endorse for the sake of escaping Kripke's puzzle. Yet we must endorse it, if Garrett's solution is to work, for it is the very fact that Pierre's two beliefs about London do *not* contradict each other,

each being qualified by a different identification of that town, that Garrett explicitly leans on to justify ascribing both beliefs to Pierre.

Robert Fogelin (1994) proposes a rather different solution to the puzzle. Fogelin argues that we should see Pierre as having what he calls a "divided belief system". Pierre's beliefs, on Fogelin's view, are divisible into two distinct subsystems, a Francophone system and an Anglophone system. Pierre's Francophone system subscribes to the belief that London is pretty; Pierre's Anglophone system subscribes to the contradictory belief. It is a mistake, on Fogelin's view, to insist on answering the question whether Pierre, considered as a whole, believes that London is pretty; we can only answer the question when it is relativized to one or the other of Pierre's two subsystems.

Some difficulties also arise for this approach to Kripke's puzzle. First, it seems to make Pierre's problem a problem of self-knowledge. If Francophone Pierre could only gain access to the beliefs of Anglophone Pierre, then perhaps he could spot the inconsistency between the two systems and make some efforts to repair it. But surely this description mistakes the case: No amount of introspective prowess can get Pierre out of his situation. What he is lacking is not some piece of knowledge about *himself,* but rather a piece of knowledge about the coreferentiality of the words 'London' and 'Londres'.

Still more troubling, to my mind, is the plethora of issues that arise about the mechanics of Fogelin's division of the mind. Francophone Pierre and Anglophone Pierre presumably share many beliefs, even if they do not agree about the aesthetic merits of London. Are these beliefs somehow encoded twice in Pierre's brain, once in English and once in French, or are they only encoded once, with the Francophone system and the Anglophone system equally capable of accessing most of them? If they are encoded twice, that seems like an awful waste of resources. If there is one common pool of beliefs to which both systems have access, how is it that beliefs, one way or the other, about London's beauty came to be excluded from that pool? What is the mechanism that separates Pierre's two subsystems of belief, and to what extent is communication possible between the parts? Fogelin also suggests other ways of dividing the mind -- for instance, a person might have beliefs in a subsystem of his mind activated when he is drunk that he does not have in the subsystem that is active when he is sober. One might ask whether different divisions of the mind can cross-cut each other; if so, can they act as a bridge for communication between those parts they cross-cut?

I put forward these questions to bring out the serious nature of the claim that the mind is divided into subsystems; claims of this sort, if they are to be taken literally, raise a variety of issues. It makes sense to consider such issues about, for example, the division of the visual system from the rest of the brain. The anatomical, neurophysiological, and cognitive

evidence for such a division is strong, and we do want to know what the mechanisms of isolation and communication are, how and whether the division cross-cuts other plausible divisions in the mind, and to what extent information must be re-encoded within different systems. It seems a radical step to say that Pierre is similarly *literally* divided into Francophone and Anglophone belief subsystems; but if the division is merely a metaphorical one, it's hard to see how it will do the necessary work.

Most people's first reaction to Kripke's puzzle is that its solution must be easy. And I do think a proper solution, which falls out of the account of belief offered in the previous chapter, has something of an easy feel about it. On the other hand, the variety and complexity of the solutions that have been offered to this puzzle belies the hunch that the problem is a cinch; we should not underplay the difficulty of Kripke's puzzle.

Self-Deception

The second philosophical puzzle I will consider is the case of self-deception. The philosophical literature on self-deception, like the literature on Kripke's puzzle, presents situations in which it is difficult to say whether a particular belief ascription is appropriate or not. Such a case is described by Amelie Rorty in a 1988 paper on the topic:

> If anyone is ever self-deceived, Dr. Laetitia Androvna is
> that person. A specialist in the diagnosis of cancer,
> whose fascination for the obscure does not usually blind
> her to the obvious, she has begun to misdescribe and
> ignore symptoms that the most junior premedical student
> would recognize as the unmistakable symptoms of the late

302

stages of a currently incurable form of cancer. Normally
introspective, given to consulting friends on important
matters, she now uncharacteristically deflects their
questions and attempts to discuss her condition.
Nevertheless, also uncharacteristically, she is bringing
her practical and financial affairs into order: Though
young and by no means affluent, she is drawing up a
detailed will. Never a serious correspondent, reticent
about matters of affection, she has taken to writing
effusive letters to distant friends and relatives,
intimating farewells, and urging them to visit her soon.
Let us suppose that none of this behavior is deliberately
deceptive: She has not adopted a policy of stoic silence
to spare her friends. On the surface of it, as far as
she knows, she is hiding nothing (1988, p. 11).

Let us now consider the following question: Does Androvna believe
that she has cancer? Different facts about Androvna seem to
point in different directions. On the one hand, Androvna's
drawing up her will and writing effusive letters are actions that
seem inexplicable unless they arise somehow from the belief that
she has cancer. On the other hand, Androvna sincerely and
consistently disavows having this belief, argues that the
evidence for cancer is inconclusive, thinks her brother rude and
ignorant when he suggests that she may have cancer, and so forth.
These actions seem difficult to explain unless we say that
Androvna does *not* believe that she has cancer. We would appear
to have, then, a dilemma: Say that Androvna does believe she has
cancer and one subset of her actions becomes inexplicable; say
that she doesn't believe it and a different subset of her actions
becomes inexplicable. Our everyday intuitions about belief
ascription don't weigh in strongly in favor of one option or the
other. The phrase 'self-deception' seems to suggest that she has
somehow managed to fool herself into believing that she doesn't

have cancer, and therefore doesn't believe that she has cancer.
On the other hand, it seems equally natural to say that certain
actions reveal that "deep down" she *does* believe that she has
cancer.

Some authors, such as Rorty (1972, 1988), David Pears (1984)
and perhaps Donald Davidson (1982a, 1985b), have attempted to
escape this dilemma by pursuing an alternative similar to
Fogelin's proposal for dealing with Kripke's puzzle: They have
suggested splintering the self into discrete subsystems, each
with only partial access to the other's cognitions. Once this is
done, the option is open to say that Androvna has one subsystem
that believes that she has cancer and another subsystem that does
not. The actions that seem to require the belief that she has
cancer are actions that are directed by, or somehow informed
especially by, the subsystem that has that belief. The actions
that seem to require absence of this belief are those directed or
informed by the other subsystem. A variant of this strategy,
advocated by Raphael Demos (1960) and Brian McLaughlin (1988),
does not strictly insist on dividing the mind into subsystems but
rather allows the unpleasant belief (in this case Androvna's
belief that she has cancer) to retreat, in some range of
circumstances, into "inaccessibility" while the contrary belief
(that she does not have cancer) is held in some more accessible
fashion.

Other authors, such as Robert Audi (1982, 1985) and Kent Bach
(1981), have argued that the self-deceiver really, genuinely

believes only the unpleasant proposition, and not the more
desirable one. Actions that seem to depend on *not* having the
unpleasant belief are then explained as the effect of suppressing
the belief or acting on the basis of persistent avowals to the
contrary. Still others, such as Alfred Mele (1987a), have argued
the opposite: What the subject really believes is rather the more
pleasant proposition -- Androvna really believes that she does
not have cancer. This belief emerges as the product of various
biasing strategies, such as weighing evidence in favor of the
preferred belief more heavily than it warrants or only making an
effort to gather evidence on one side of the issue. If one
occasionally acts, as Androvna does, on the basis of the
unpleasant truth, doing so must be the product of a momentary
lapse in one's ordinarily more pleasant convictions.

Each of these approaches to self-deception has some
plausibility, and it is difficult to find a firm basis on which
to choose between them -- although I mentioned in my discussion
of Fogelin some reasons I have to be hesitant about strategies
like Rorty's and Pears' that involve partitioning the mind. But
if we cannot easily choose between these accounts, neither can we
endorse all of them, since they are incompatible.

The Puzzles Resolved

I think that the cases of both Pierre and Androvna are cases
of in-between believing. It is a mistake to insist on a definite
resolution to the question of whether Pierre or Androvna really

305

have the beliefs that intuition ambivalently attributes and denies to them. And once we let go of the inclination to insist on simple answers to questions about what they believe, the puzzles disappear.

There are actually two steps involved in this approach to the puzzles. The first step is to reject the original Simple Question formulations of the puzzles -- that is, refuse to answer Kripke's insistent question about whether Pierre really does or really does not believe that London is pretty, and, likewise, to refuse to answer the question of whether Androvna really does or really does not believe that she has cancer. So far, the move is not a new one. Both Garrett and Fogelin agree that the question, "Does Pierre believe that London is pretty?" cannot, as it stands, get a simple yes-or-no answer. This point is also argued by Laurence Goldstein (1993) and Graeme Forbes (1994). In the self-deception literature the option of refusing to say that either "yes the self-deceived person believes the unpleasant proposition" or "no she doesn't" is surprisingly uncommon. One sees this view, perhaps, in H. O. Mounce's (1971) paper on the subject, and Mele describes it as an option in a review article on self-deception (1987b), although he neither accepts the idea nor specifically addresses it in his positive work on the topic (1987a).

The more original element of my approach comes with the second step in the resolution of these puzzles. One wants not only to make the negative move just described, but also to develop a *positive* description of the cases at hand. Although

mere recognition of the existence of in-between states of believing may be sufficient to suggest the rejection of the Simple Question in the cases of Pierre and Androvna, a more positive vision of the nature of belief must guide the attempt to give a full satisfactory account of these cases. Here is where my approach diverges from that of Garrett and Fogelin, despite our agreement about the need to reject Kripke's Simple Question.

The difference is that Garrett and Fogelin both allow an all-or-nothing view of belief to re-enter through the back door. Fogelin, although he refuses to say that Pierre, considered as a whole person, either believes or does not believe that London is pretty, does think that Pierre is divisible into *parts* for which simple yes-or-no answers to these questions are appropriate. Similarly Garrett, although he refuses to say that Pierre either believes or does not believe the unqualified proposition that London is pretty, does think that Pierre fully and completely believes the proposition that London, identified in the way associated with the name 'Londres', is pretty, and that Pierre fully and completely believes the proposition that London, identified in the way associated with the name 'London', is not pretty. Both Fogelin and Garrett, then, seem to be seeking some way of carving up affairs so that all legitimate questions about belief *can* get simple yes or no answers. They simply reject the idea that Kripke's original question about Pierre is a legitimate question.

The approach I recommend for describing cases such as Pierre's and Androvna's is the same approach I recommend for describing the multifarious variety of other in-between cases of believing. We should describe the dispositional make-up of the subject at hand, looking both at behavioral dispositions and at phenomenal dispositions; and then we should *stop*. We may, if we wish, note which dispositional patterns match up with which belief stereotypes; we may inquire as to how the subject came to have such a mixed set of dispositions, or how the subject might bring herself better into line with the stereotypes. But these are questions that stand apart from the question of what the subject believes.

There is something approximately right in describing Pierre as believing that London is pretty and in describing Androvna as believing that she has cancer. Both Pierre and Androvna have a number of dispositions that accord with these beliefs, and describing them as having these beliefs can be pragmatically workable to the extent that we can focus our attention and interest on *these* dispositions and explain away with plausible mechanisms other dispositions that accord less well with the stereotypes. At the same time, and for the same reasons, there is something approximately right in describing Pierre as believing that London is not pretty and in describing Androvna as believing that she does not have cancer. But the only completely accurate answer to the question of what Pierre and Androvna believe is an answer that conveys the full mix of their

dispositions without attempting to squeeze them into any of the stereotypes.

Philosophers and psychologists may have felt it necessary to force in-between cases of believing into a simple yes-or-no paradigm because there has been no good picture of belief enabling them to do otherwise. The maneuvers of Kripke, Fogelin, Rorty, and others might then be seen in a Kuhnian (1970) light, as attempts to deal with anomalous data or problem cases by pushing them into the best existing paradigms. My hope is that by presenting a dispositional account of belief and discussing its relation to in-between cases of believing, I have made plausible the claim that there *is* a good alternative to insisting that the only real answers to questions about belief must be of the yes-or-no (or possibly the "degree of belief") variety. Describing a subject as having a divergence of dispositions on a topic is, on my view, *not* settling for less than a full answer to the question of what she believes.

2. What's in a Look?

Major revolutions in a child's cognitive development, like major revolutions in science, do not typically take place all in an instant, but are, as I have repeatedly emphasized, gradual and protracted affairs. If these revolutions can be characterized as changes in (among other things) the child's beliefs, they should be an abundant source of examples of the kind of "in-between" beliefs that are the focus of these chapters. One should positively *expect* periods of in-between believing. I will examine here two cases in which developmental psychologists have been led astray by the inclination to regard the child's knowledge in an all-or-nothing manner. I will begin by exploring Renée Baillargeon's influential views on the infant's understanding of the existence of unperceived objects, and then I will turn to some recent work by Wendy Clements and Josef Perner on the child's understanding of false belief.

The Child's Understanding of Object Permanence

Renée Baillargeon is interested in discovering at what age the child comes to understand that an object observed at two distinct moments in time must also exist in the period between observations. Her work on this topic (e.g., Baillargeon 1987; Baillargeon et al. 1985; Baillargeon and DeVos 1991; Baillargeon, et al. 1990) grows out of a tradition beginning with Piaget (1954). Piaget regards the acquisition of this knowledge about objects as crucial in the development of the concept of "object

permanence", which he sees progressing through several stages between roughly the ages of six and eighteen months. Piaget observed that if a toy in which an infant is interested is removed from view by being placed, in full view of the infant, under a blanket or behind an occluder, children under nine months will not search for it, even though they may have the motor ability to lift blankets and peek behind occluders. It is as though, for the infant, the object no longer existed. Gopnik and Meltzoff (1996), and Harris (1983, 1987) provide interesting reviews of the extensive literature on the development of the object concept.

I will take some time to describe Baillargeon's best-known experiment designed to test the infant's knowledge of object permanence (1987). I will then describe her conclusions from this experiment and provide some arguments against them.

The experimental subjects, 3 1/2- and 4 1/2-month-old infants, were first allowed to handle and were thus familiarized with a 25 x 15 x 5 cm. yellow wooden box with a clown face on it. The infant was then placed before a platform on which a large silver screen lay flat and the yellow box was visible standing upright behind it. The box was then removed and the infant entered the "habituation phase" of the experiment.

In the habituation phase, the large silver screen before the infant was slowly rotated back and forth several times through 180° of arc. The screen began flat on the platform, its top facing the infant, was slowly raised 90° to an upright position, and then was slowly lowered to lay flat against on the platform,

facing away from the infant, having completed 180° of arc. The screen then reversed its path, coming up through 90° and at the end of the cycle lying flat with its top again toward the infant. One cycle took approximately 10 seconds.

The habituation phase acquainted the infant with the motion of the screen and provided a measure against which the infants' looking times at the control and the test events could be measured. A "habituation trial" consisted of a series of cycles, terminating when the infant either (a.) looked away for 2 consecutive seconds after having looked at the display for at least 5 cumulative seconds or (b.) looked at the event for 60 cumulative seconds. Habituation trials were repeated until the infant's looking time on three consecutive trials was 50% or less than her average looking time on the first three trials or until nine cycles were completed, whichever came first.

The infants were then divided into experimental and control conditions. In the experimental condition, the infants were shown two different events, an "impossible event" and a "possible event", in an alternating sequence, until each event had been observed four times. Half the infants saw the impossible event first, and half saw the possible event first.

The "impossible event" began with the screen lying flat toward the infant and the yellow box visible on the platform behind it. The screen was then rotated through 180° of arc, as in the habituation event, while the yellow box was surreptitiously removed so that it would not interfere with the motion of the screen through its last degrees of arc. After

completing the 180º, the screen would reverse its path and the yellow box would be surreptitiously replaced so that at the end of the event the box would be visible again and the screen flat toward the infant. The cycle was then repeated. The end of a trial was determined by the same criteria as the end of a habituation trial. These trials were dubbed the "impossible event" trials because they convey (to an adult) the impression of the screen "impossibly" passing through or squeezing flat the yellow box during its last degrees of arc.

The "possible event" was like the impossible event, except that the screen only rotated through 112º of arc, stopping before hitting the yellow box. The screen then reversed its path to lie flat before the infant with the box visible behind it.

The control conditions were like the experimental conditions, except that the box was absent. Infants in the control condition watched four alternating pairs of 180º and 112º events, just as the infants in the experimental conditions did.

The diagram below, which illustrates some aspects of the conditions just described, is a modified version of a diagram presented in Baillargeon (1987).
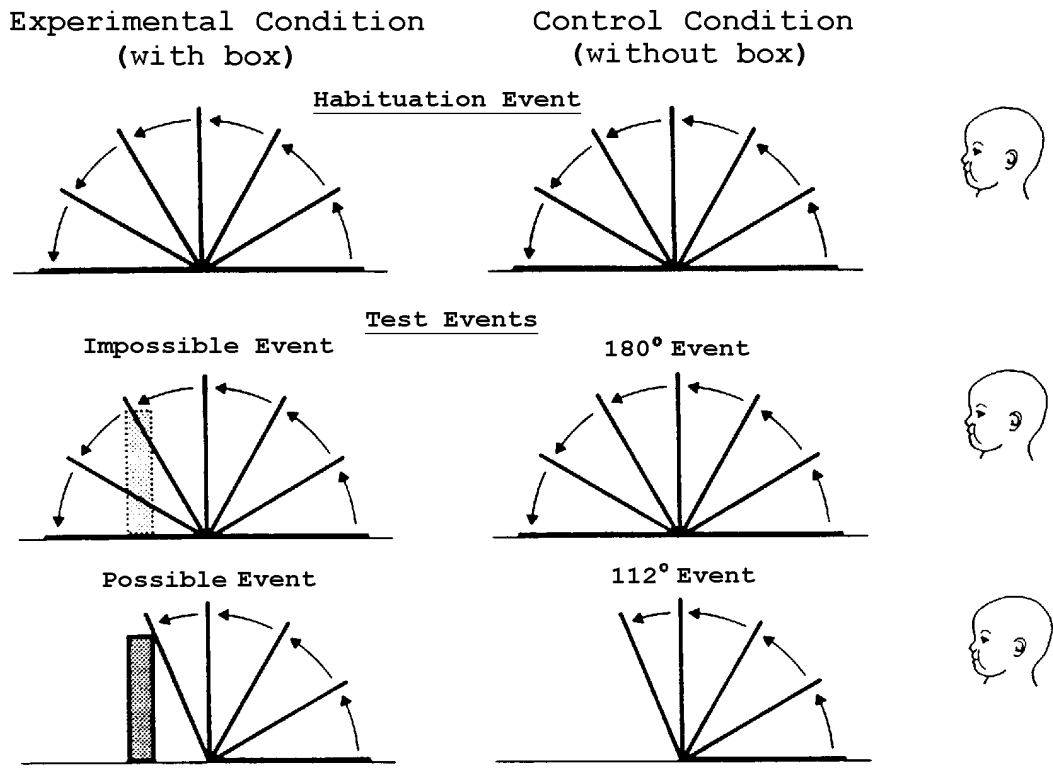
Experimental Condition (with box)   Control Condition (without box)

Habituation Event

Test Events

Impossible Event   180° Event

Possible Event   112° Event

Figure 1. Schematic representation of the habituation and test events shown to the infants in the experimental and control conditions in Baillargeon (1987).

Baillargeon's interest was in the looking times of the infants in the test conditions. This study, as do all "habituation studies", relies on the presupposition that infants will look longer at events that differ more from the event to which they were "habituated" than at those that differ less from the habituation event. Although one could raise methodological questions about this assumption, that is not my plan here. In accordance with the habituation assumption, Baillargeon anticipated that if an infant looked longer at the impossible test events than at the possible ones, that would be because the

314

infants saw the former as differing more from the habituation

event than the latter. This in turn, Baillargeon argues, could

only be explained if we assumed that the infant knew that the

objects continued to exist even when they were not being

perceived. In superficial respects, the 180° impossible event is

more like the 180° control event than the 112° possible event is;

it is only if one takes into account the apparent "squeezing" or

removal of the occluded box that the 180° event seems strange or

unique.

Baillargeon found that 4 1/2-month-olds (and "fast

habituating" 3 1/2-month-olds) did look significantly longer at

the impossible event than at the possible one. This increase in

looking time cannot be explained simply by the infants'

preferring to watch the screen rotate through 180° over watching

it rotate through 112°, because infants in the control condition

did not exhibit such a preference. It is natural to suppose,

then, that the infants looked longer at the impossible event than

the possible one because it violated their expectations about the

world.

Baillargeon concludes that, contra Piaget, "infants as young

as 3 1/2 months of age already realize that objects continue to

exist when occluded" (1987, p. 662). At the same time, she does

not deny Piaget's claim that infants' search activities do not

reveal such an understanding until the period between nine and

eighteen months of age.

The developmental difference between a four-month-old and a

nine-month-old is dramatic. The question then arises: If the

infant *really* understands object permanence at four months, why
is this understanding not revealed in the child's searching
behavior until nine months at least? Baillargeon recognizes this
as a difficulty and indicates that the problem may be with the
child's means-ends reasoning -- her ability to apply an action to
one object (e.g., pull a blanket) to create conditions in which
she may apply another action to a different object (e.g., grab
the hidden toy). Nevertheless, Baillargeon refers to Piaget's
(1952) observations of sequences of behavior in three- and four-
month-olds in which an action is applied to one object (e.g., a
chain) to produce an effect in another object (e.g., shaking a
toy attached to the other end of the chain). Why the latter kind
of means-ends reasoning should be available so early and the
former kind so late, and what differentiates the two, Baillargeon
admits to be "somewhat of a mystery" (1987, p. 663).

With the latter remarks, Baillargeon may be making things
harder for herself than she needs to: Piaget doesn't claim really
to find means-ends reasoning involving distinct objects until
around nine months of age -- the same age at which he discovers
search behavior revealing some knowledge of object permanence.
The three- or four-month old who pulls a chain to shake an object
at the end of it may not see the objects at hand as a system of
two separate objects causally related to each other. On the
other hand, by six or seven months a child who will not remove an
obstacle wholly occluding a desired object will move an obstacle
partly occluding the desired object (Piaget 1954) and will move a

transparent cover wholly enclosing the desired object (Bower and Wishart 1972), so the problem cannot *simply* be with moving one object to get to another.

In fact, studies on infants in this age range yield starkly divided results on the question of whether infants can reason about objects outside their perceptual fields. On the one hand, when the lights are turned off or the infant rotates her head away from an object, she seems to keep track of its existence (Piaget 1954; Bower and Wishart 1972; Clifton, Rochat, Litovsky, and Perris 1991). Young infants are also able to track the motion of an object as it passes behind an occluder and anticipate the point of reappearance on the other side, sometimes even looking back to the point of disappearance if the object does not reappear in the predicted location (Bower, Broughton, and Moore 1971). There have also been a number of other studies suggesting that infants dishabituate to or look preferentially at events seeming to require that, while occluded, either one object has passed through another or an object has taken a discontinuous path, or in which the number of revealed objects after a period of occlusion is different than an adult would anticipate in the circumstances (Baillargeon 1991; Baillargeon et al. 1985; Spelke et al. 1992; Spelke et al. 1994; Moore et al. 1978; Wynn 1992). On the other hand, a number of studies provide evidence against the idea that infants truly understand that objects continue to exist unperceived. Not only do Piaget's (1954) observations on reaching suggest this, but so also do studies showing that

infants under nine months do not seem disturbed when an object
disappears behind one edge of a screen and reappears at the other
without making any appearance crossing a gap in the middle of the
screen (Moore et al. 1978; but see Baillargeon and Devos 1991);
nor do infants show anticipatory reaching for objects on occluded
trajectories, though they will reach for objects on visible
trajectories (von Hofsten 1994, cited in Gopnik and Meltzoff
1997).

Many of the psychologists conducting such studies, not least
Baillargeon, seem committed to arguing one way or another
regarding the question of whether the infant genuinely believes
that objects continue to exist unperceived. Results pointing in
the other direction must then be either discredited or left
mysterious.

This area of development would seem to be an excellent
example of one on which we ought to say that the infants *neither*
truly believe that objects continue to exist unperceived nor
truly fail to understand this. Instead, their dispositions on
the matter are mixed. Shouldn't one expect infants, in the
course of gradual development, to pass through a period like this
in any case? It may be time for us to stop beating ourselves
over the head looking for a simple yes-or-no answer to the
question of whether six-month-olds have an understanding of
object permanence. A more useful project would be to determine
exactly which of their dispositions point in which directions,

318

how things came to be this way, and how they change over the course of time.

Implicit Understanding of False Belief?

I conclude with a second case from the developmental literature, Clements' and Perner's paper, "Implicit Understanding of Belief" (1994).

Clements and Perner tested children from 2 1/2 years to 4 1/2 years of age on a variation of the classic false belief experiments performed by Wimmer and Perner (1983) and reported in chapter two. In Clement's and Perner's study, children were introduced to Sam Mouse and his two mouse holes, connected by a V-shaped tunnel. In front of one mouse hole was a red box; in front of the other was a blue box. Children were told the following story, which was simultaneously enacted with cardboard cutouts:

> This is Sam. One day Sam had some cheese for tea. When he looked there was one piece of cheese left but he was too full up to eat it. "I know," he said. "I'll put it in this blue box and I can eat it later." Sam gave a big yawn. "I'm so tired now," he said. He went all the way down the tunnel and went to bed where he fell fast asleep (Clements and Perner 1994, p. 382).

After checking that the child remembered the location of the cheese, the story was continued.

> When Sam had fallen fast asleep, Katie came back from playing outside. As she walked past the blue box, she looked into it and saw the cheese. "Oh look!" she said. "Someone's left a piece of cheese here. I'll put it in the red box and I can eat it later for my tea." So she picked up the cheese and walked, fully visible, across the hill to the other mouse hole where she put the cheese

> in the red box. "I'll go and see my friend now" she said (Clements and Perner 1994, p. 382).

After asking several questions assuring that the child remembered important facts about the plot, the story was brought to its dramatic conclusion.

> Later on, Sam woke up and gave a big stretch. "I feel very hungry now," he said. "I'll go and get the cheese" (Clements and Perner 1994, p. 383).

The experimenter then said, "I wonder where he's going to look?" and paused for one or two seconds for the child to think about where Sam would look. Throughout this time, the child's eye movements were recorded on videotape. Finally, the experimenter reminded the child that "Sam wants to get the cheese" and concluded by asking the child two questions: "Which box will he open?" and "Why do you think he will open that box?"

A control group heard much the same story, only with Sam watching while Katie moved the cheese. Half the children heard one of these stories starring Sam Mouse, while half of the children heard a similar story starring Sarah, whose letter was carried by the wind from the upper to the lower balcony of her house.

Children over four tended both to look at the correct box in response to the experimenter's prompt "I wonder where he's going to look" and to answer the false belief question correctly. Children under two years, eleven months did exactly the opposite. The interesting results in this study were from the children in the middle age range, from two years eleven months to three years eleven months. Children in this age range typically answered the

false belief question incorrectly but *looked* at the correct box in response to the experimenter's prompt question.

What could explain these results? Clements and Perner reject the hypothesis that in the false belief condition the children are looking at the box in which Sam first left his cheese simply because they are retracing the events of the story, in light of the fact that they don't do this in the control condition and instead look directly at the correct box.[1] Another possibility Clements and Perner reject is that the children's looking reflects tentative hypotheses they momentarily entertain. If this were the case, Clements and Perner argue, the children should have looked at least as frequently at the box they ultimately (and mistakenly) claimed Sam would open as at the other box. Instead, the children look *consistently* at the correct box.

Clements and Perner think the child's eye motions in this experiment reflect some genuine anticipation of Sam's looking in the box in which he originally placed the cheese. Supposing we grant them this, something of a puzzle arises. If the child really understands that Sam will look where he originally left the cheese, why does the child *say* that Sam will look in the other box? Alternately, if the child really *doesn't* understand that Sam will look in the wrong place, how can her eye movements

---

[1] One possibility Clements and Perner do not consider is that the children retraced the story with their eyes *only* when asked the confusing false belief question. Even such a possibility, however, requires that the children were alert enough to false beliefs that they found the false belief task confusing and the control task simple. Thus, it may still reflect the "implicit" understanding of false belief Clements and Perner argue children this age have.

correctly anticipate the place he will look? The reader may by now scent the likely presence of a mixed dispositional profile.

Clements and Perner could, of course, escape their dilemma by rejecting the Simple Question, but they do not. Instead, they argue that "the eye movements reveal a different type of knowledge" than that revealed by the verbal responses (p. 391). In particular, the eye movements reveal "implicit" knowledge, the verbal responses "explicit" knowledge. Clements and Perner also characterize the difference as one between "nonjudgmental" and "judgmental" knowledge (p. 392). They explain further:

> That is, pure action (i.e. looking in anticipation) is done only on the basis of a representation of reality; that is, one model. But to make a judgment (verbally or gesturally) at least two models are required: One to represent the proposition to be judged (information expressed), and the other to represent the state of the world by which this proposition is to be judged. In other words, to make a judgment is to convey that the verbally or otherwise expressed information (the model of whatever is being proposed) conforms with reality (the other model) (p. 392-393).

Following Karmiloff-Smith (1992), they generalize:

> So, whenever knowledge is acquired in a new domain (acquired procedurally or through abstraction of observed regularities), it becomes first available nonjudgmentally before it can be used to make judgments. For that reason, children in our study are able to anticipate the protagonist's movements correctly with their eyes before they can make a judgment about where he will go (p. 393).

The idea, then, is that major developmental changes in knowledge may be generally first reflected in nonverbal, "nonjudgmental" behavior and only later realized in verbal judgments. And why might this be so? Because verbal judgments are more complicated than "pure action": Pure action requires only that the subject have a correct indicative representation of the world, while

322

verbal judgment requires combining this with an assessment of the truth or falsity of a proposition expressing these facts about the world.

I am actually sympathetic to the idea that some nonverbal dispositions may be acquired before verbal ones belonging to the same stereotype (although one might imagine this pattern reversed in the case of things taught at school), but Clement's and Perner's view, like many built on all-or-nothing assumptions about belief, adds needless machinery to this observation. What divides "pure actions" not requiring assessment of a proposition from actions like speech that do require such an assessment (and thus two "models") remains something of a mystery. Would opening the correct box instead of naming or pointing to it be a "pure action"? What about interfering with Sam's journey there? The distinction between judgmental and nonjudgmental knowledge must inherit the blurriness of the distinction between pure actions and judgmental ones.

Another problem with Clement's and Perner's view reveals itself as well. Whatever the line between judgmental and nonjudgmental knowledge, conscious verbal assessments must belong to the former category. But even at the same age we see the anticipatory looking, if that's what it is, other signs of judgmental knowledge of false belief are emerging in a limited range of contexts, such as when the child is asked to explain mistaken actions *after* they have occurred (Wellman 1990), and when the child is specifically engaged in the task of "tricking"

323

someone (Sullivan and Winner 1993). Even, then, if we granted

that the distinction between judgmental and nonjudgmental

knowledge was a clear one, it would not be motivated by the false

belief literature. The picture we see is instead that of a child

slowly acquiring the knowledge of false belief: In her early

threes, a very few of her dispositions accord with this

knowledge, and as she ages, more and more of her dispositions do.

It does no good to attempt to salvage all-or-nothing intuitions

about belief with the claim that the three-year-old really, fully

has one species of knowledge and really, fully lacks another

species. The facts are simply not so clean as that.

3. Conclusion

The last three chapters have been occupied with the motivation, explanation, and defense of a novel account of belief, what I have called the *phenomenal dispositional* account of belief. This account arises from the need for an approach to belief that can make sense of *in-between* cases of believing, cases in which the subject is not accurately describable with the everyday "yes-or-no" patterns of belief ascription. The account treats believing as nothing more or less than having dispositions that match stereotypical dispositional profiles. Cases of in-between believing are then treated as cases in which the subject fails to match cleanly with any stereotypical dispositional belief profile.

Several debates in the philosophical and developmental literatures were discussed with the tools provided by the dispositional account, and were shown to profit from the use of those tools. Of particular importance was the ability of the dispositional account to focus its subscriber's interest on problems *other* than trying to extract a simple yes-or-no answer to the question of whether a subject whose dispositional profile is mixed has a particular belief. Trying to force in-between cases of believing into an all-or-nothing mold not only imposes a misleading simplicity on these cases, but also raises a tricky dilemma: On the one hand, if the subject really does fully and completely have the belief, how it is possible that she does not manifest it in a wide variety of circumstances? On the other

hand, if the subject does not really have the belief, how can it be that she seems sometimes to act on the basis of the knowledge denied her? It is tempting to try to escape this dilemma by inventing mental machinery, as we have seen in the cases of Fogelin, Rorty, and Clements and Perner. The danger in this move is not that dividing the mind or introducing different faculties of believing is in itself a mistake, but rather that its postulation in these cases is only as justified as the resolution to describe these cases in an all-or-nothing manner.

Besides the danger of insisting too adamantly on discovering simple yes-or-no answers to questions about what a subject believes, however, is the converse danger -- that of giving up too quickly in finding such answers. In chapter five I outlined a primary reason for seeking such yes-or-no answers: People generally conform fairly well to the stereotypes, and evidence pointing toward a mixed dispositional profile will often sort itself out clearly in favor of one stereotype or another. It is important to distinguish cases in which a person only *seems* to have mixed dispositions from cases in which the nonconformity is genuine. Good judgment will have to be our guide in deciding when to concede the presence of a genuinely mixed dispositional profile and thus to give up on finding simple yes-or-no answers to what the subject believes. The judgment is complicated by the presence of more than simply epistemic factors. The yes-or-no approach also has the advantage of simplicity, which may in some contexts outweigh the increased accuracy of more detailed

dispositional descriptions when the subject fairly closely

matches one stereotype or another; and furthermore, insistence on

simple yes-or-no questions about belief may also serve the

purpose of motivating both ourselves and others to conform to

societially necessary dispositional stereotypes, as suggested in

chapter six.

Besides having these reasons for insisting on yes-or-no

answers to questions about belief, philosophers and psychologists

may have felt it necessary to force in-between cases of believing

into a simple yes-or-no paradigm because there has been no good

scientific alternative allowing one to do otherwise. The

maneuvers of Kripke, Fogelin, Baillargeon, and the others might

then be seen generously, in a Kuhnian (1970) light, as attempts

to deal with anomalous data or problem cases by pushing them into

the best existing paradigm. One could hardly expect a good

scientist to do otherwise. My hope is that these chapters have

convinced the reader that there *is* a good scientific alternative

to insisting that the only "real" answers to questions about

belief must be of the yes-or-no (or possibly the "degree of

belief") variety -- and that describing a subject as having a

divergence of dispositions on a topic is not settling for less

and provides no hindrance to scientific research. It is worth

noting in this regard that neural net models of cognition

(classically described in Rumelhart et al. 1986 and McClelland et

al. 1986) seem to allow quite naturally a broad range of in-

between responses and dispositional mixes, and that if neural net

models find broad use in understanding human cognition, an account of cognition that can handle these "in-betweenish" features of neural nets will be necessary.

So much for the pragmatic benefits of the account. The ontological dimension of the account is, I hope, conservative and widely acceptable. There is, of course, some talk about properties, dispositions, and stereotypes, but I do not believe it has been necessary to take any controversial stands on these matters. I have claimed that in-between cases of believing are common, and I have provided a number of examples of such cases. While any individual example may itself be controversial as an instance of in-between believing, what is important to my position in not any individual case, but rather the overall impression I sought to create of the ubiquity of such in-between cases.

One ontological claim, however, is crucial to my account and at the same time potentially controversial. It is the claim that once one has fully described a subject's dispositional profile and compared that profile to the relevant stereotypes, one has exhausted everything we can know about what that subject believes on the topic. There is no *further* fact of the matter, apart from facts about the subject's dispositional profile, about what the subject "really" believes -- or at least no fact we can presently discover. It is unclear what would count as a discovery of such a fact (unless we consider the possibility of science eventually developing in the direction suggested at the end of chapter six),

and accounts like the one offered here show, I think, that we can run philosophy and the sciences without appeal to such facts. Occam's razor, then, recommends leaving them out of our ontology.

One can view the project of this chapter as a revamping of the old Rylean dispositionalist view of belief, with a new emphasis on the phenomenal aspects of the account. This project is quite timely in its way. The 1990's have seen a resurgence of academic interest in the phenomenal aspects of mind (Searle 1992 is an excellent example), and I should not be surprised to see quite a number of mid-century views reincarnated with a phenomenal twist.