Chapter 4—
The Epistemological Problem

The epistemological problem has two halves, both of them concerned with how we come to have *knowledge* of the internal activities of conscious, intelligent minds. The first problem is called the *problem of other minds*: How does one determine whether something other than oneself—an alien creature, a sophisticated robot, a socially active computer, or even another human—is really a thinking, feeling, conscious being; rather than, for example, an unconscious automaton whose behavior arises from something other than genuine mental states? How can one tell? The second problem is called the *problem of self-consciousness*: How is it that any conscious being has an immediate and privileged knowledge of its own sensations, emotions, beliefs, desires, and so forth? How is this possible? And just how trustworthy is that knowledge? The solutions to these problems, I think it will emerge, are not independent. Let us explore the first problem first.

1—
The Problem of Other Minds

It is of course by observing a creature's behavior, including its verbal behavior, that we judge it to be a conscious, thinking creature—to be 'another mind'. From bodily damage and moaning, we infer pain. From smiles and laughter, we infer joy. From the dodging of a snowball, we infer perception. From complex and appropriate manipulation of the environment, we infer desires, intentions, and beliefs. From these and other things, and above all from speech, we infer conscious intelligence in the creature at issue.

This much is obvious, but these remarks serve only to introduce our problem, not to solve it. The problem begins to emerge when we ask what *justifies* the sorts of inferences cited. To infer the (hidden) occurrence of certain kinds of mental states from the occurrence of certain kinds of behavior is to assume that appropriate general connections hold between them, connections presumably of the form, "If behavior of kind *B* is displayed by any creature, then usually a mental state of

kind *S* is occurring." Such 'psychobehavioral generalizations' have the form of standard empirical generalizations, such as "If a sound like thunder occurs, then usually there is (or was) a lightning strike somewhere in the vicinity." Presumably their justification is parallel also: such general statements are justified by our past experience of a regular connection between the phenomena cited. Wherever and whenever we perceive lightning, we generally perceive (very loud) thunder also, and barring the machinery of war, nothing else produces exactly that sound.

But how can one be justified in believing that the relevant psychobehavioral generalizations are true of other creatures, *when all one can ever observe is one-half of the alleged connection: the creature's behavior*? The creature's mental states, if he happens to have any, are directly observable only by the creature himself. We cannot observe them. And so we cannot possibly gather the sort of empirical support needed. Apparently then, one cannot possibly be justified in believing in such psychobehavioral generalizations. Accordingly, one cannot be justified in drawing inferences from another creature's behavior, to his possession of mental states. Which is to say, one cannot be justified in believing that any creature other than oneself has mental states!

This conclusion is deeply implausible, but the skeptical problem is quite robust. Belief in other minds requires inferences from behavior; such inferences require generalizations about creatures in general; such generalizations can only be justified by experience of creatures in general; but experience of one's own case is all one can have. This is the classical problem of other minds.

The Argument from Analogy

There are three classical attempts at a solution to the problem of other minds, and perhaps the simplest of these is the *argument from analogy*. One can observe both halves of the psychobehavioral connections in exactly one case, it is argued: in one's own. And one can determine that the relevant generalizations are indeed true, at least of oneself. But other humans are, so far as I am able to observe, entirely similar to me. If the generalizations are true of me, then it is a reasonable inference, by analogy with my own case, that they are also true of other humans. Therefore, I do have some justification for accepting those generalizations after all, and I am therefore justified in drawing specific inferences about the mental states of specific creatures on the strength of them.

Our impulse to resist the skeptical conclusion of the problem of other minds is sufficiently great that we are likely to grasp at any solution that promises a way around it. There are serious difficulties with the argument from analogy, however, and we should be wary of accepting

it. The first problem is that it represents one's knowledge of other minds as resting on an inductive generalization from exactly *one* case. This is absolutely the weakest possible instance of an inductive argument, comparable to inferring that all bears are white on the strength of observing a single bear (a polar bear). It may well be wondered whether our robust confidence in the existence of other minds can possibly be accounted for and exhausted by such a feeble argument. Surely, one wants to object, my belief that you are conscious is better founded than *that*.

And there are further problems. If one's knowledge of other minds is ultimately limited by what one can observe in one's own case, then it will not be possible for color-blind people justly to believe that other humans have visual sensations that are denied to them, nor possible for a deaf person justly to believe that others can hear, and so forth. One can reasonably ascribe to other minds, on this view, only what one finds in one's own mind. This entails, for example, that one could not possibly be justified in ascribing mental states to an alien creature, if its psychology were systematically different from one's own (as, after all, it is likely to be). Are one's reasonable hypotheses about the contents of other minds really limited in these parochial ways?

A third objection attempts to undercut the argument from analogy entirely, as an account of how we come to appreciate the psychobehavioral connections at issue. If I am to distinguish between and clearly recognize the many varieties of mental states, thereafter to divine the connections they bear to my behavior, I must possess the concepts necessary for making such identifying judgments: I must grasp the meaning of the terms "pain", "grief", "fear", "desire", "belief", and so forth. But we have already seen from the preceding chapter that the meaning of those terms is given, largely or entirely, by a network of general assumptions connecting them with terms for other mental states, external circumstances, and observable behavior. Simply to possess the relevant concepts, therefore, is *already* to be apprised of the general connections between mental states and behavior that the examination of one's own case was supposed to provide. One's understanding of our folk-psychological concepts, therefore, must derive from something more than just the uninformed examination of one's own stream of consciousness.

Collectively, these difficulties with the argument from analogy have provided a strong motive for seeking a different solution to the problem of other minds. One that does not create problems of the same order as the problem to be solved.

Behaviorism Again

The philosophical behaviorists were quick to press a different solution, one informed by the difficulties discovered in the argument from analogy. Specifically, they argued that if the generalizations connecting mental states with behavior cannot be suitably justified by empirical observation, then perhaps that is because those generalizations were not empirical generalizations to start with. Rather, it was suggested, those generalizations are true by sheer *definition*. They are operational definitions of the psychological terms they contain. As such, they stand in no need of empirical justification. And a creature that behaves, or is disposed to behave, in the appropriate ways *is by definition* conscious, sentient, and intelligent. (Typical behaviorists were not always this bold and forthright in their claims, but neither were they often this clear in what they claimed.)

Given the pressure to solve the other-minds problem, the impotence of the argument from analogy, and the appeal of the idea that the meaning of psychological terms was in some way bound up with psychobehavioral generalizations, one can appreciate why philosophers tried so hard to make some variant of this position work. But they failed. When we examine the generalizations of folk psychology, we find that they seldom if ever take the form of simple 'operational definitions' (recall the discussion of the term "soluble" in 2.2). Behaviorists were unable to state the necessary and sufficient *behavioral* conditions for the application of even a single psychological term. Neither do the generalizations of folk psychology appear to be true by definition. They seem rather to be rough empirical truths, both in their effect on our linguistic intuitions and in their explanatory and predictive functions in everyday commerce. This fact returns us to the problem of trying to *justify* the various psychobehavioral generalizations on which one's knowledge of other minds seems to depend.

Explanatory Hypotheses and Folk Psychology

The problem of other minds was first formulated at a time when our grasp of the nature of theoretical justification was still rather primitive. Until fairly recently almost everybody believed that a general law could be justified only by an inductive generalization from a suitable number of observed instances of the elements comprehended by the law. One sees a number of crows, notices that each of them is black, and one generalizes to "All crows are black". And so for any law. It was thought. This idea might have been adequate for laws connecting observable things and properties, but modern science is full of laws governing the behavior of *unobservable* things and properties. Think of atoms, and molecules, and genes, and

electromagnetic waves. Plainly, laws concerning unobservables must enjoy some other form of empirical justification, if they are to be justified at all.

This other form of justification is not far to seek. Theorists postulate unobservable entities, and specific laws governing them, because occasionally this produces a theory that allows us to construct predictions and explanations of observable phenomena hitherto unexplained. More specifically, if we assume certain hypotheses, and conjoin with them information about observable circumstances, we can often deduce statements concerning further observable phenomena, statements which, it subsequently turns out, are systematically *true*. To the degree that any theory displays such explanatory and predictive virtues, that theory becomes a beliefworthy. hypothesis. It has what is commonly called "hypothetico-deductive" justification (or "H-D" justification, for short). In sum, a theory about unobservables can be beliefworthy if it allows us to explain and to predict some domain of observable phenomena better than any competing theory. This is in fact the standard mode of justification for theories in general.

Consider now the network of general principles—connecting mental states with one another, with bodily circumstances, and with behavior—that constitutes folk psychology. This 'theory' allows us to explain and to predict the behavior of human beings better than any other hypothesis currently available, and what better reason can there be for believing a set of general laws about unobservable states and properties? The laws of folk psychology are beliefworthy for the same reason that the laws of any theory are beliefworthy: their explanatory and predictive success. Note further that one's justification here need owe nothing at all to one's examination of one's own case. It is folk psychology's success with respect to the behavior of people in general that matters. Conceivably one's own case might even differ from that of others (recall the 'alien creature' objection to the argument from analogy). But this need not affect one's theoretical access to their internal states, however different they might be. One would simply use a different psychological theory to understand their behavior, a theory different from the one that comprehends one's own inner life and outer behavior.

Turning now from general laws to individuals, the hypothesis that a specific individual has conscious intelligence is also an explanatory hypothesis, on this view. And it is plausible to the degree that the individual's continuing behavior is best explained and predicted in terms of desires, beliefs, perceptions, emotions, and so on. Since that is, in fact, the best way to understand the behavior of most humans, one is therefore justified in believing that they are 'other minds'. And one will be similarly justified in ascribing psychological states to any

other creatures or machines, so long as such ascriptions sustain the most successful explanations and predictions of their continuing behavior.

Thus the most recent solution to the problem of other minds. Its virtues are fairly straightforward, and it coheres nicely with our earlier solution to the semantical problem. Both problems seem to yield to the assumption that our common-sense conceptual framework for mental states has all the features of a theory. Not everyone has found this assumption plausible, however, its virtues notwithstanding. If you center your attention on your direct consciousness of your own mental states, the idea that they are 'theoretical entities' may seem a very strange suggestion. Whether and how that suggestion might make sense will be addressed in the next section.

Suggested Readings

Malcolm, Norman, "Knowledge of Other Minds," *Journal of Philosophy*, vol. LV (1958). Reprinted in *The Philosophy of Mind*, ed. V. C. Chappell (Englewood Cliffs, NJ: Prentice-Hall, 1962).

Strawson, Sir Peter, "Persons," in *Minnesota Studies in the Philosophy of Science*, vol. II eds. H. Feigl, M. Scriven, and G. Maxwell (Minneapolis: University of Minnesota Press, 1958). Reprinted in *The Philosophy of Mind*, ed. V. C. Chappell (Englewood Cliffs, NJ: Prentice-Hall, 1962).

Sellars, Wilfrid, "Empiricism and the Philosophy of Mind," in *Minnesota Studies in the Philosophy of Science*, vol. I, eds. H. Feigl and M. Scriven (Minneapolis: University of Minnesota Press, 1956). Reprinted in Wilfrid Sellars, *Science, Perception, and Reality* (London: Routledge & Keegan Paul, 1963), sections 45–63.

Churchland, Paul, *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge University Press, 1979), section 12.

2—
The Problem of Self-Consciousness

Upon first reflection, self-consciousness is likely to seem implacably mysterious and utterly unique. This is part of what makes it so fascinating. Upon deeper reflection, however, the veil of mystery begins to part just a little, and self-consciousness can be seen as one instance of a more general phenomenon.

To be self-conscious is to have, at a minimum, *knowledge* of oneself. But this is not all. Self-consciousness involves knowledge not just of one's physical states, but knowledge of one's *mental states* specifically. Additionally, self-consciousness involves the same kind of *continuously updated* knowledge that one enjoys in one's continuous perception of the external world. Self-consciousness, it seems, is a kind of continuous apprehension of an inner reality, the reality of one's mental states and activities.

Self-Consciousness:
A Contemporary View

The point about apprehension is important: evidently it is not enough just to have mental states. One must discriminate one kind of state from another. One must recognize them for what they are. In sum, one must apprehend them within some conceptual framework or other that catalogs the various different types of mental states. Only then will recognitional *judgment* be possible ("I am angry", "I am elated", "I believe that *P*", and so on). This suggests that there are different degrees of self-consciousness, since presumably one's ability to discriminate subtly different types of mental states improves with practice and increasing experience, and since the conceptual framework within which explicit recognition is expressed grows in sophistication and comprehensiveness as one learns more and more about the intricacies of human nature. Accordingly, the self-awareness of a young child, though real, will be much narrower or coarser than that of a sensitive adult. What is simply a dislike of someone, for a child, may divide into a mixture of jealousy, fear, and moral disapproval of someone, in the case of an honest and self-perceptive adult.

This suggests further that self-consciousness may vary from person to person, depending on which areas of discrimination and conception have been most thoroughly mastered. A novelist or psychologist may have a running awareness of her emotional states that is far more penetrating than the rest of us enjoy; a logician may have a more detailed consciousness of the continuing evolution of his beliefs; a decision-theorist may have a superior awareness of the flux of her desires, intentions, and practical reasonings; a painter may have a keener recognition of the structure of his visual sensations; and so forth. Self-consciousness, evidently, has a very large *learned* component.

In these respects, one's introspective consciousness of oneself appears very similar to one's perceptual consciousness of the external world. The difference is that, in the former case, whatever mechanisms of discrimination are at work are keyed to internal circumstances instead of to external ones. The mechanisms themselves are presumably innate, but one must learn to use them: to make useful discriminations and to prompt insightful judgments. Learned perceptual skills are familiar in the case of external perception. A symphony conductor can hear the clarinets' contribution to what is a seamless sound to a child. An astronomer can recognize the planets, and nebulae, and red giants, among what are just specks in the night sky to others. A skilled chef can taste the rosemary and shallots within what is just a yummy taste to a hungry diner. And so forth. It is evident that perception, whether inner or outer, is substantially a learned skill. Most of that learning takes place in our early childhood, of course: what is perceptually obvious to us now was a subtle discrimination at eight months. But there is always room to learn more.

In summary, self-consciousness, on this view, is just a species of perception: *self-perception*. It is not perception of one's foot with one's eyes, for example, but is rather the perception of one's internal states with what we may call (largely in ignorance) one's faculty of introspection. Self-consciousness is thus no more (and no less) mysterious than perception generally. It is just directed internally rather than externally.

Nor is it at all surprising that cognitively advanced creatures should possess self-consciousness. What perception requires is no more than that one's faculty of judgment be in systematic causal contact with the domain to be perceived, in such a way that we can learn to make, on a continuing basis, spontaneous, noninferred, but appropriate judgments about that domain. Our faculty of judgment is in causal contact with the external world, through the various sensory modalities; but it is also in systematic causal contact with the rest of the internal domain of which it is a part. Who will express surprise that one kind of brain activity enjoys rich causal connections with other kinds of brain activity? But such connections carry information, and thus make 'informed' judgment possible. Self-consciousness, therefore, at some level or to some degree of penetration, is to be expected in almost any cognitively advanced creature.

This view coheres with an evolutionary view. Presumably humanity has struggled toward self-consciousness in two dimensions: in the neurophysiological evolution of our ability to make useful introspective discriminations, and in the social evolution of a conceptual framework to exploit that discriminative capacity in prompting explanatorily and

predictively useful judgments. As well, each of us is the site of an evolutionary struggle towards self-apprehension during his or her lifetime, in which we learn to use and to refine the innate discriminative capacities, and to master the socially entrenched conceptual framework (folk psychology) necessary to exploit them.

The Traditional View

These remarks on self-consciousness may seem plausible enough, but a long tradition in the philosophy of mind takes a very different view of our introspective knowledge. Introspection, it has been argued, is fundamentally different from any form of external perception. Our perception of the external world is always mediated by sensations or impressions of some kind, and the external world is thus known only indirectly and problematically. With introspection, however, our knowledge is immediate and direct. One does not introspectively apprehend a sensation by way of a sensation of that sensation, or apprehend an impression by way of an impression of that impression. As a result, one cannot be the victim of a false impression (of an impression), or a misleading sensation (of a sensation). Therefore, once one is considering the states of one's own mind, the distinction between appearance and reality disappears entirely. The mind is transparent to itself, and things in the mind are, necessarily, exactly what they 'seem' to be. It does not make any sense to say, for example, "It seemed to me that I was in considerable pain, but I was mistaken." Accordingly, one's candid introspective judgments about one's own mental states—or about one's own *sensations*, anyway—are incorrigible and infallible: it is logically impossible that they be mistaken. The mind knows itself first, in a unique way, and far better than it can ever know the external world.

This extraordinary position must be taken seriously—at least temporarily—for several reasons. First, it is part and parcel of an old and influential theory of knowledge-in-general: orthodox empiricism. Second, the claim that one's knowledge of one's sensations is unmediated, by further 'sensations$_2$', does seem plausible. And any attempt to deny it would lead either to an infinite regress of 'sensations$_3$', 'sensations$_4$', and so on; or to some level of 'sensations$_n$', where one's knowledge of them *is* at last unmediated. Third, the proponent of this view has a powerful rhetorical question. "How *could* one possibly be mistaken about whether or not one is in *pain*? How is it even possible to be wrong about a thing like that?" As the reader will note, this question is not easy to answer.

Arguments against the Traditional View

The view that the mind knows itself first, in a unique way, and far better than it can ever know the external world, has dominated Western thought for over three centuries. But if one adopts a thoroughly naturalistic and evolutionary perspective on the mind, the traditional view quickly acquires a sort of fairy-tale quality. After all, brains were selected for because brains conferred a reproductive advantage on the individuals that possessed them. And they conferred that advantage because they allowed the individuals to anticipate their environment, to distinguish food from nonfood, predators from nonpredators, safety from peril, and mates from nonmates. In sum, a brain gave them knowledge and control of *the external world*. Brains have been the beneficiaries of natural selection precisely because of that feature. Evidently, what they know first and best is not themselves, but the environment in which they have to survive.

The capacity for *self*-knowledge could conceivably be selected for as the incidental concomitant of the capacity for knowledge generally, and it might be selected for specifically if it happened to enhance in some way the brain's capacity for external knowledge. But in either case it would be at best a secondary advantage, derivative upon the increase in one's knowledge and control of the external world. And in any case, there is no reason to assume that self-perception, to the extent that it did evolve, would be fundamentally different in kind from external perception; and no reason at all to assume that it would be infallible.

If the traditional view is basically implausible, let us examine the arguments set forth in its favor, and see whether they withstand scrutiny. Consider first the rhetorical question, "How could one possibly be mistaken about the identity of one's own sensations?" As an argument for the incorrigibility of our knowledge of our sensations, it has the form, "None of us can *think* of a way in which we could be mistaken in our judgments about our sensations; therefore there *is* no way in which we could be mistaken." But this commits an elementary fallacy: it is an argument from ignorance. There may well be ways in which error is possible, despite our ignorance of them. Indeed, perhaps we are unaware of them precisely because we understand so little about the hidden mechanisms of introspection. The rhetorical question, therefore, could safely be put aside, even if we could not answer it. But in fact we can. With a little effort, we can think of many ways in which errors of introspective judgment can and do occur, as we shall see presently.

Consider now the argument that the distinction between appearance and reality must collapse in the case of sensations, since our apprehension of them is not mediated by anything that might misrepresent

them. This argument is good only if misrepresentation by an intermediary is the only way in which errors could occur. But it is not. Even if introspection is unmediated by second-order 'sensations$_2$', nothing guarantees that the introspective judgment, "I am in pain," will be caused only by the occurrence of pains. Perhaps other things as well can cause that judgment, at least in unusual circumstances, in which case the judgment would be false. Consider the occurrence of something rather *similar* to pain—a sudden sensation of extreme cold, for example—in a situation where one strongly *expects* to feel pain. Suppose you are a captured spy, being interrogated at length with the repeated help of a hot iron pressed briefly to your back. If, on the twentieth trial, an *ice cube* is covertly pressed against your back, your immediate reaction will differ little or none from your first nineteen reactions. You almost certainly would think, for a brief moment, that you were feeling pain.

The incorrigibilist may try to insist that sensation number twenty was a pain after all, despite its benign cause, on the grounds that, if you took it to be a pain, if you thought it felt painful to you, then it really was a pain. This interpretation sits poorly with the fact that one can recover from the kinds of misidentification just explored. One's initial screech of horror gives way to "Wait . . . wait . . . that's not the same feeling as before. What's going on back there??" If sensation number twenty really *was* a pain, why does one's judgment reverse itself a few seconds later?

A similar case: the taste-sensation of lime sherbet is only very slightly different from the taste-sensation of orange sherbet, and in blindfold tests people do surprisingly poorly at telling which sensation is which. An orange-expectant subject fed lime sherbet may confidently identify her taste-sensation as being of the kind normally produced by orange sherbet, only to retract the identification immediately upon being given a (blind) taste of the genuinely orange article. Here one *corrects* one's qualitative identification, in flat contradiction to the idea that mistakes are impossible. Mistakes of this kind are called *expectation effects*, and they are a standard phenomenon with perception generally. Evidently, they apply to introspection as well. The reality of expectation effects provides us with a recipe for producing almost any misidentification you like, whether of external things or of internal states.

Further, do we really know enough about the mechanisms of introspection to insist that nothing mediates the sensation and the judgment about it? Granted, there is no intermediary that we are *aware* of, but this means nothing, since on any view there must be much of the mind's operation that is below the level of introspective detection. Here then is another possible source of error. The distinction between ap-

pearance and reality may be hard to draw, in the case of sensations, only because we know so little about the ways in which things can and do go wrong.

Another way in which sensations can be misjudged emerges when we consider sensations with very short durations. Sensations can be artificially induced so as to have durations of arbitrary length. Not surprisingly, as the durations become shorter, reliable identifications (of their qualitative identity) become harder and harder to make, and mistakes become, not impossible, but inevitable. Which is to say, the agreement between what the subject says the sensation is, and what its mode of production indicates it should be, is near-perfect for long presentations, but falls off toward chance as the length of the presentations approaches zero. Such 'presentation effects' are also standard in perception generally. And if the subject is suitably drugged or exhausted, the reliability of his identifications falls off even more swiftly. This too is standard.

Memory effects must also be mentioned. Suppose a person who, perhaps because of some neural damage in youth, has not felt pain or any other tactile or visceral sensation for *fifty years*, or has been colorblind for the same period. Does anyone really suppose that, if the subject's neural deficit were suddenly repaired after such a long hiatus, he would instantly be able to discriminate and identify (= recognize which similarity-class is instanced) every one of his newly recovered sensations, and do it with infallible accuracy? The idea is not at all plausible. Similar effects could also be produced in the short term, with a drug that temporarily clouds one's memory of the various types of sensations. Failures of identification and outright misidentifications would then be wholly natural. And even in the normal case, are spontaneous, isolated, and unnoticed lapses of memory utterly impossible? How can the defender of the traditional view rule them out?

A more familiar sort of case also merits mention. Suppose you are dreaming that you have a splitting headache, or that you are in excruciating pain from being tortured. When you awaken suddenly, do you not realize, in a wave of relief, that you were *not really* the victim of a headache, or of excruciating pain, despite the conviction that attends every dream? The incorrigibility thesis is beginning to look highly implausible.

None of this should be surprising. The incorrigibility thesis might have been initially plausible in the case of sensations, but it is not remotely plausible for most other mental states like beliefs, desires, and emotions. We are notoriously bad, for example, at judging whether we are jealous, or vindictive; at judging our most basic desires; and at judging our own traits of character. Granted, infallibility has seldom

been claimed for anything beyond sensations. But this restriction raises problems of its own. Why should infallibility attend sensations, but not emotions and desires? Knowledge of the latter seems no more 'mediated' than knowledge of the former.

Intriguingly, recent research in social psychology has shown that the explanations one offers for one's own behavior often have little or no origin in reliable introspection, despite one's sincere beliefs to that effect, but are instead spontaneously confabulated on the spot as *explanatory hypotheses* to fit the behavior and circumstances observed (see the paper by Nisbett and Wilson cited in the suggested readings at the end of this section). And they are often demonstrably wrong, since the 'introspective' reports given prove to be a function of wholly external features of the experimental situation, features under the control of the experimenters. On the view of these researchers, much of what passes for introspective reports is really the expression of one's spontaneous *theorizing* about one's reasons, motives, and perceptions, where the hypotheses produced are based on the same external evidence available to the public at large.

Consider a final argument against the incorrigibility thesis. Our introspective judgments are framed in the concepts of folk psychology, which framework we have already determined (in chapters 3.3, 3.4, and 4.1) to have the structure and status of an empirical theory. As with any such judgments, their integrity is only as good as the integrity of the empirical theory in which the relevant concepts are semantically embedded. Which is to say, if folk psychology should turn out to be a radically false theory, then its entire ontology would lose its claim to reality. And any judgment framed in its terms would have to be deemed false by reason of presupposing a false background theory. Since folk psychology is an empirical theory, it is always strictly possible that it might turn out to be radically false. Accordingly, it is always possible that any judgment framed in its terms be false. Therefore, our introspective judgments are not incorrigible. Not only might they be wrong occasionally, and one by one; they might *all* be cockeyed!

The Theory-Ladenness of All Perception

The strangeness of the idea that mental states are 'theoretical' can be reduced by the following reflections. *All* perceptual judgments, not just introspective ones, are 'theory-laden': all perception involves speculative interpretation. This, at least, is the claim of more recently developed versions of empiricism. The basic idea behind this claim can be expressed with the following very brief, but very general, argument: the *network argument*.

1. Any perceptual judgment involves the application of *concepts* (for example, *a* is *F*).

2. Any concept is a node in a *network* of contrasting concepts, and its meaning is fixed by its peculiar place within that network.

3. Any network of concepts is a speculative assumption or *theory*: minimally, as to the classes into which nature divides herself, and the major relations that hold between them.

Therefore,

4. Any perceptual judgment presupposes a theory.

According to this general view, the mind/brain is a furiously active theorizer from the word go. The perceptual world is largely an unintelligible confusion to a newborn infant, but its mind/brain sets about immediately to formulate a conceptual framework with which to apprehend, to explain, and to anticipate that world. Thus ensues a sequence of conceptual inventions, modifications, and revolutions that finally produces something approximating our common-sense conception of the world. The furious conceptual evolution undergone by every child in its first two years is probably never equaled throughout the remainder of its life.

The point of all this, for our purposes, is as follows. At life's opening, the mind/brain finds itself as confusing and unintelligible as it finds the external world. It must set about to learn the structure and activities of its inner states no less than it must set about to learn the structure and activities of the external world. With time, it does learn about itself, but through a process of conceptual development and learned discrimination that parallels exactly the process by which it apprehends the world outside of it. The traditional view, it would seem, is simply mistaken.

Suggested Readings

Armstrong, David, *A Materialist Theory of the Mind* (London: Routledge & Keegan Paul, 1968), chapter 6, sections IX, X; and chapter 15, section II.

Dennett, Daniel, "Toward a Cognitive Theory of Consciousness," in *Minnesota Studies in the Philosophy of Science*, vol. IX, ed. C. W. Savage (Minneapolis: University of Minnesota Press, 1978). Reprinted in Daniel Dennett, *Brainstorms* (Montgomery, VT: Bradford, 1978; Cambridge, MA: MIT Press).

Nisbett, Richard, and Wilson, Timothy, "Telling More Than We Can Know: Verbal Reports on Mental Processes," *Psychological Review*, vol. 84, no. 3 (1977).

Churchland, Patricia, "Consciousness: The Transmutation of a Concept," *Pacific Philosophical Quarterly*, vol. 64 (1983).

Churchland, Paul, *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge

University Press, 1979), sections 13 and 16; on the theory-ladenness of perception in general, see chapter 2.

Nagel, Thomas, "What Is It Like to Be a Bat?" *Philosophical Review*, vol. LXXXIII (1974). Reprinted in *Readings in Philosophy of Psychology*, vol. I, ed. N. Block (Cambridge, MA: Harvard University Press, 1980).