Philosophy and the Mirror of Nature

RICHARD RORTY

Princeton University Press
Princeton, New Jersey

CHAPTER II

Persons Without Minds

1. The Antipodeans

Far away, on the other side of our galaxy, there was a planet on which lived beings like ourselves—featherless bipeds who built houses and bombs, and wrote poems and computer programs. These beings did not know that they had minds. They had notions like "wanting to" and "intending to" and "believing that" and "feeling terrible" and "feeling marvelous." But they had no notion that these signified *mental* states—states of a peculiar and distinct sort—quite different from "sitting down," "having a cold," and "being sexually aroused." Although they used the notions of believing and knowing and wanting and being moody of their pets and their robots as well as of themselves, they did not regard pets or robots as included in what was meant when they said, "We all believe . . ." or "We never do such things as. . . ." That is to say, they treated only members of their own species as *persons*. But they did not explain the difference between persons and non-persons by such notions as "mind," "consciousness," "spirit," or anything of the sort. They did not *explain* it at all; they just treated it as the difference between "us" and everything else. They believed in immortality for themselves, and a few believed that this would be shared by the pets or the robots, or both. But this immortality did not involve the notion of a "soul" which separated from the body. It was a straightforward matter of bodily resurrection followed by mysterious and instantaneous motion to what they referred to as "a place above the heavens" for good people, and to a sort of cave, beneath the planet's surface, for the wicked. Their philosophers were concerned primarily with four topics: the nature of Being, proofs of the existence of a Benevolent and

Omnipotent Being who would carry out arrangements for the resurrection, problems arising out of discourse about nonexistent objects, and the reconciliation of conflicting moral intuitions. But these philosophers had not formulated the problem of subject and object, nor that of mind and matter. There was a tradition of Pyrrhonian skepticism, but Locke's "veil of ideas" was unknown, since the notion of an "idea" or "perception" or "mental representation" was also unknown. Some of their philosophers predicted that the beliefs about immortality which had been central in earlier periods of history, and which were still held by all but the intelligentsia, would someday be replaced by a "positivistic" culture purged of all superstitions (but these philosophers made no mention of an intervening "metaphysical" stage).

In most respects, then, the language, life, technology, and philosophy of this race were much like ours. But there was one important difference. Neurology and biochemistry had been the first disciplines in which technological breakthroughs had been achieved, and a large part of the conversation of these people concerned the state of their nerves. When their infants veered toward hot stoves, mothers cried out, "He'll stimulate his C-fibers." When people were given clever visual illusions to look at, they said, "How odd! It makes neuronic bundle G-14 quiver, but when I look at it from the side I can see that it's not a red rectangle at all." Their knowledge of physiology was such that each well-formed sentence in the language which anybody bothered to form could easily be correlated with a readily identifiable neural state. This state occurred whenever someone uttered, or was tempted to utter, or heard, the sentence. This state also sometimes occurred in solitude and people reported such occasions with remarks like "I was suddenly in state S-296, so I put out the milk bottles." Sometimes they would say things like "It looked like an elephant, but then it struck me that elephants don't occur on this continent, so I realized that it must be a mastodon." But they would also

sometimes say, in just the same circumstances, things like "I had G-412 together with F-11, but then I had S-147, so I realized that it must be a mastodon." They thought of mastodons and milk bottles as objects of beliefs and desires, and as causing certain neural processes. They viewed these neural processes as interacting causally with beliefs and desires—in just the same way as the mastodons and milk bottles did. Certain neural processes could be deliberately self-induced, and some people were more skillful than others in inducing certain neural states in themselves. Others were skilled at detecting certain special states which most people could not recognize in themselves.

In the middle of the twenty-first century, an expedition from Earth landed on this planet. The expedition included philosophers, as well as representatives of every other learned discipline. The philosophers thought that the most interesting thing about the natives was their lack of the concept of mind. They joked among themselves that they had landed among a bunch of materialists, and suggested the name Antipodea for the planet—in reference to an almost forgotten school of philosophers, centering in Australia and New Zealand, who in the previous century had attempted one of the many futile revolts against Cartesian dualism in the history of Terran philosophy. The name stuck, and so the new race of intelligent beings came to be known as Antipodeans. The Terran neurologists and biochemists were fascinated by the wealth of knowledge in their field which the Antipodeans exhibited. Since technical conversation on these subjects was conducted almost entirely in offhand references to neural states, the Terran experts eventually picked up the ability to report their own neural states (without conscious inference) instead of reporting their thoughts, perceptions, and raw feels. (The physiologies of the two species were, fortunately, almost identical.) Everything went swimmingly, except for the difficulties met by the philosophers.

The philosophers who had come on the expedition were,

as usual, divided into two warring camps: the tender-minded ones who thought philosophy should aim at Significance, and the tough-minded philosophers who thought that it should aim at Truth. The philosophers of the first sort felt that there was no real problem about whether the Antipodeans had minds. They held that what was important in understanding other beings was a grasp of their mode of being-in-the-world. It became evident that, whatever *Existentiale* the Antipodeans were using, they certainly did not include any of those which, a century earlier, Heidegger had criticized as "subjectivist." The whole notion of "the epistemological subject," or the person as spirit, had no place in their self-descriptions, nor in their philosophies. Some of the tender-minded philosophers felt that this showed that the Antipodeans had not yet broken out of Nature into Spirit, or, more charitably, had not yet progressed from Consciousness to Self-Consciousness. These philosophers became town-criers of inwardness, attempting to bully the Antipodeans across an invisible line and into the Realm of Spirit. Others, however, felt that the Antipodeans exhibited the praiseworthy grasp of the union of πόλεμος and λόγος which was lost to Western Terran consciousness through Plato's assimilation of οὐσία to ἰδέα. The Antipodean failure to grasp the notion of mind, in the view of this set of philosophers, showed their closeness to Being and their freedom from the temptations to which Terran thought had long since succumbed. In the contest between these two views, equally tender-minded as both were, discussion tended to be inconclusive. The Antipodeans themselves were not much help, because they had so much trouble translating the background reading necessary to appreciate the problem—Plato's *Theaetetus*, Descartes's *Meditations*, Hume's *Treatise*, Kant's *Critique of Pure Reason*, Hegel's *Phenomenology*, Strawson's *Individuals*, etc.

The tough-minded philosophers, as usual, found a much more straightforward and clean-cut question to discuss. They did not care what the Antipodeans thought about

themselves, but rather focused on the question: Do they in fact have minds? In their precise way, they narrowed this question down to: Do they in fact have sensations? It was thought that if it became clear whether they had, say, sensations of pain, as well as stimulated C-fibers, when touching hot stoves, everything else would be plain sailing. It was clear that the Antipodeans had the same behavioral dispositions toward hot stoves, muscle cramps, torture, and the like as humans. They loathed having their C-fibers stimulated. But the tough-minded philosophers asked themselves: Does their experience contain the same phenomenal properties as ours? Does the stimulation of C-fibers feel painful? Or does it feel some other, equally awful, way? Or does feeling not come into it at all? These philosophers were not surprised that the Antipodeans could offer noninferential reports of their own neural states, since it had been learned long since that psychophysiologists could train human subjects to report alpha-rhythms, as well as various other physiologically describable cortical states. But they felt baffled by the question: Are some phenomenal properties being detected by an Antipodean who says, "It's my C-fibers again—you know, the ones that go off every time you get burned or hit or have a tooth pulled. It's just awful."?

It was suggested that the question could only be answered experimentally, and so they arranged with the neurologists that one of their number should be wired up to an Antipodean volunteer so as to switch currents back and forth between various regions of the two brains. This, it was thought, would also enable the philosophers to insure that the Antipodeans did not have an inverted spectrum, or anything else which might confuse the issue. As it turned out, however, the experiment produced no interesting results. The difficulty was that when the Antipodean speech center got an input from the C-fibers of the Earthling brain it always talked only about its C-fibers, whereas when the Earthling speech center was in control it always talked only about pain. When the Antipodean speech center was asked

what the C-fibers felt like it said that it didn't quite get the notion of "feeling," but that stimulated C-fibers were, of course, terrible things to have. The same sort of thing happened for the questions about inverted spectra and other perceptual qualities. When asked to call off the colors on a chart, both speech centers called off the usual color-names in the same order. But the Antipodean speech center could also call off the various neuronic bundles activated by each patch on the chart (no matter which visual cortex it happened to be hooked up to). When the Earthling speech center was asked what the colors were like when transmitted to the Antipodean visual cortex, it said that they seemed just as usual.

This experiment seemed not to have helped. For it was still obscure whether the Antipodeans had pains. It was equally obscure whether they had one or two raw feels when indigo light streamed onto their retinas (one of indigo, and one of neural state C-692)—or whether they had no raw feels at all. The Antipodeans were repeatedly questioned about how they knew it was indigo. They replied that they could see that it was. When asked how they knew they were in C-692, they said they "just knew" it. When it was suggested to them that they might have unconsciously inferred that it was indigo on the basis of the C-692 feel, they seemed unable to understand what unconscious inference was, or what "feels" were. When it was suggested to them that they might have made the same inference to the fact that they were in state C-692 on the basis of the raw feel of indigo, they were, of course, equally baffled. When they were asked whether the neural state appeared indigo, they replied that it did not—the *light* was indigo—and that the questioner must be making some sort of category mistake. When they were asked whether they could imagine having C-692 and not seeing indigo, they said they could not. When asked whether it was a conceptual truth or an empirical generalization that these two experiences went together, they replied that they were not sure how to tell the difference. When asked

whether they could be wrong about whether they were seeing indigo, they replied that they of course could, but could not be wrong about whether they seemed to be seeing indigo. When asked whether they could be wrong about whether they were in state C-692, they replied in exactly the same way. Finally, skillful philosophical dialectic brought them to realize that what they could not imagine was seeming to see indigo and failing to seem to be in state C-692. But this result did not seem to help with the questions: "Raw feels?" "Two raw feels or one?" "Two referents or one referent under two descriptions?" Nor did any of this help with the question about the way in which stimulated C-fibers appeared to them. When they were asked whether they could be mistaken in thinking that their C-fibers were stimulated, they replied that of course they could—but that they could not imagine being mistaken about whether their C-fibers seemed to be stimulated.

At this point, it occurred to someone to ask whether they could detect the neural state which was the concomitant of "seeming to have their C-fibers stimulated." Antipodeans replied that there was, of course, the state T-435 which was the constant neural concomitant of the utterance of the sentence "My C-fibers seem to be stimulated," state T-497 which went with "It's just as if my C-fibers were being stimulated," state T-293 which went with "Stimulated C-fibers!" and various other neural states which were concomitants of various other roughly synonymous sentences—but that there was no further neural state which they were aware of in addition to these. Cases in which Antipodeans had T-435 but no stimulation of C-fibers included those in which, for example, they were strapped to what they were falsely informed was a torture machine, a switch was theatrically turned on, but nothing else was done.

Discussion among the philosophers now switched to the topic: Could the Antipodeans be mistaken about the T-series of neural states (the ones which were concomitants of understanding or uttering sentences)? Could they seem to

be having T-435 but not really be? Yes, the Antipodeans said, cerebroscopes indicated that sort of thing occasionally happened. Was there any explanation of the cases in which it happened—any pattern to them? No, there did not seem to be. It was just one of those odd things that turned up occasionally. Neurophysiology had not yet been able to find another sort of neural state, outside the T-series, which was a concomitant of such weird illusions, any more than for certain perceptual illusions, but perhaps it would someday.

This answer left the philosophers still in difficulties on the question of whether the Antipodeans had sensations of pain, or anything else. For there now seemed to be nothing which the Antipodeans were incorrigible about except how things seemed to them. But it was not clear that "how things seemed to them" was a matter of what raw feels they had, as opposed to what they were inclined to say. If they had the raw feel of painfulness, then they had minds. But a raw feel is (or has) a phenomenal property—one which you cannot have the illusion of having (because, so to speak, having the illusion of it is itself to have *it*). The difference between stimulated C-fibers and pains was that you could have the illusion of stimulated C-fibers (could have, e.g., T-435) without having stimulated C-fibers, but could not have the illusion of pain without having pain. There was nothing which the Antipodeans could not be wrong about except how things seemed to them. But the fact that they could not "merely seem to have it seem to them that . . ." was of no interest in determining whether they had minds. The fact that "seems to seem . . ." is an expression without a use is a fact about the notion of "appearance," not a tip-off to the presence of "phenomenal properties." For the appearance-reality distinction is not based on a distinction between subjective representations and objective states of affairs; it is merely a matter of getting something wrong, having a false belief. So the Antipodeans' firm grasp of the former distinction did not help philosophers tell whether to ascribe the latter to them.

2. Phenomenal Properties

Coming back now to the present, what *should* we say about the Antipodeans? The first thing to do, presumably, is to look more closely at the notion of "phenomenal property," and in particular at the disanalogy between apprehending a physical phenomenon in a misleading way and apprehending a mental phenomenon in a misleading way. Kripke's account of the distinction sums up the intuition on which defenders of dualism have usually relied, so we may begin a closer look by trying to apply his terminology:

> Someone can be in the same epistemic situation as he would be if there were heat, even in the absence of heat, simply by feeling the sensation of heat; and even in the presence of heat, he can have the same evidence as he would have in the absence of heat simply by lacking the sensation S. No such possibility exists in the case of pain or in other mental phenomena. To be in the same epistemic situation that would obtain if one had a pain *is* to have a pain; to be in the same epistemic situation that would obtain in the absence of a pain *is* not to have a pain. . . . The trouble is that the notion of an epistemic situation qualitatively identical to one in which the observer had a sensation *S* simply *is* one in which the observer had that sensation. The same point can be made in terms of the notion of what picks out the reference of a rigid designator [an expression which designates the same object in all the possible worlds in which it designates at all]. In the case of identity of heat with molecular motion the important consideration was that although "heat" is a rigid designator, the reference of that designator was determined by an accidental property of the referent, namely the property of producing in us the sensation *S*. . . . Pain, on the other hand, is not picked out by one of its accidental properties; rather it is picked out by the property of being pain itself, by its immediate phenomenological quality. Thus pain, unlike heat, is not only rigidly designated by "pain" but the reference of the designator is determined by an essential property of the referent. Thus it is not possible to say that although pain is necessarily identical with a certain physical state, a certain phenomenon can be picked out in the same way we pick out pain without being correlated with that physical state. If any phenomenon is picked out in exactly the same way that we pick out pain, then that phenomenon *is* pain.[1]

These considerations suggest that the real question is: Do the Antipodeans pick out mental phenomena by accidental properties? If we assume for the moment that they *do* have pains, could they perhaps miss the "immediate phenomenological quality" and note only the accidental feature of being constantly accompanied by stimulated C-fibers? Or, if they cannot exactly *miss* an immediate phenomenological quality, might they perhaps fail to have a name for it, and thus fail to pick out the entity that has the quality by an essential property? To put it another way, since the Antipodeans do *not* pick out pain "in exactly the same way that we pick out pain," can we conclude that whatever they have it is *not* pain? Is one's epistemic relation to one's raw feels necessary as well as sufficient to establish the existence of the raw feel in question? Or should we say that actually they *do* pick out pain in exactly the way that we do—because when they say, "Ooh! Stimulated C-fibers!" they feel exactly what we feel when we say, "Pain!"? Actually, perhaps, they were feeling pain and calling that feeling "the state of seeming to have one's C-fibers stimulated," and they are in the same epistemic situation relative to seeming to have their C-fibers stimulated as we are in seeming to see something red, and to all other such incorrigible states.

[1] Saul Kripke, "Naming and Necessity" in *Semantics of Natural Language*, ed. Donald Davidson and Gilbert Harman (Dordrecht, 1972), pp. 339-340. For criticism of Kripke's discussion of dualism and materialism, see Fred Feldman, "Kripke on the Identity Theory," and William Lycan, "Kripke and the Materialists," both in *Journal of Philosophy* 71 (1974), 665-689.

It now looks as if what we need is some quite general criterion for deciding when two things are "really" the same thing described in two different ways. For there seems nothing distinctive about the present conundrum which makes it depend upon the peculiarities of the mental. If we agree that what counts in deciding whether the Antipodeans have raw feels is incorrigibility—the inability to have an illusion of . . .—the general problem about alternative descriptions will still prevent us from applying this criterion and thus resolving the issue. This problem is not one which is going to receive a neat, clear-cut, readily applicable solution. For nothing general will resolve every tension between saying,

> You're talking about X's all right, but practically everything you say about them is false

and saying instead,

> Since practically nothing you say is true of X's, you can't be talking about X's.

But let us put aside this difficulty for the moment (returning to it in chapter six) and consider the still more depressing point that anyone who even tried to state general criteria for assimilating or distinguishing referents of expressions would need some general ontological categories—some firm, if coarse, way of blocking things out—just to get started. It would help, in particular, to have a distinction between mental entities and physical entities. But the problem about the Antipodeans puts this whole distinction in doubt. To see why it does so, suppose that there are no criteria for "mental phenomenon" save Kripke's epistemic one.[2] This supposition identifies "the mental" with raw feels, passing thoughts, and mental images. It excludes such

[2] I defended a qualified form of this supposition in "Incorrigibility as the Mark of the Mental," *Journal of Philosophy* 67 (1970), 399-424. See also Jaegwon Kim, "Materialism and the Criteria of the Mental," *Synthese* 22 (1972), 323-345, esp. 336-341.

things as beliefs, moods, and the like (which, though indubitably "higher," are nonetheless not parts of our incorrigibly reportable inner life, and hence not such as to encourage the Cartesian kind of distinction between two ontological realms). The supposal amounts, in other words, to the claim that (1) it is sufficient for being a mental state that the thing in question be incorrigibly knowable by its possessor, and (2) we do not literally attribute any nonphysical states (e.g., beliefs) to beings which fail to have some such incorrigibly knowable states. (This conforms to Antipodean practice, as well as to our intuition that dogs have nonphysical states simply by virtue of having pains, whereas computers do not, even by virtue of offering us novel and exciting truths.) On this supposition, then, there will be *nothing* to answer to the question "When they report that their C-fibers seem to be firing, are they reporting a feeling (perhaps the same feeling that we report by "pain!") or are they just making the noises which are triggered by their neurons being in certain states?" And if this is so, since the role played in our lives by reports of feelings is the same as the role played in Antipodean lives by reports of neurons, we face the further question: Are *we* reporting feelings or neurons when *we* use "pain"?

To see that this is a real issue, consider the implications of the identity of functional role. If it is the case that the Antipodeans have the entire range of culture that we do, if they are as intentional in their discourse and as self-consciously aesthetic in their choice of objects and persons as we, if their yearning for moral excellence and immortality is as great, they are likely to think our philosophers' interest in whether they have minds is a bit parochial. Why, they wonder, does it make such a difference? Why, they may ask us, do *we* think that we have these odd things called "feels" and "minds"? Now that they have taught us micro-neurology, cannot we see that talk of mental states was merely a place-holder for talk of neurons? Or, if we really do have some funny extra states besides the neurological ones, are

they really all that important? Is the possession of such states really the basis for a distinction between ontological categories?

These last sets of questions illustrate how lightly the Antipodeans take the controversy which, among Terran philosophers, is the hard-fought issue between materialists and epiphenomenalists. Further, the success of Antipodean neurology, not only in the explanation and control of behavior but in supplying the vocabulary for the Antipodean self-image, shows that none of the other Terran theories about "the relation between mind and body" can even get a look-in. For parallelism and epiphenomenalism can only be differentiated on some non-Humean view of causation—some view according to which there is a causal mechanism to be discovered which will show which way causal lines run. But nobody, not even the most diehard Cartesian, imagines that when a molecule-by-molecule account of the neurons is before us (as, *ex hypothesi*, it is before the Antipodeans) there will still be a place to look for further causal mechanisms. (What would "looking" amount to?) So even if we abandon Hume, we are still in no position to be parallelist, except on some a priori ground according to which we "just know" that the mental is a self-contained causal realm. As for interactionism, the Antipodeans would not dream of denying that beliefs and desires, for example, interact causally with irradiations of the retina, movements of the arm, and so on. But they view talk of such an interaction not as yoking different ontological realms but as a handy (because brief) reference to function rather than to structure. (It is as philosophically unproblematic as a transaction between a government and an individual. No set of necessary and sufficient conditions stated in terms of just who did what to whom can be given for a remark about such a transaction, any more than for remarks about beliefs caused by radiations and movements caused by beliefs—but who would have thought they could?) Interaction would only be of interest if a neural discharge were swerved

from its course by a raw feel, or drained of some of its power by a raw feel, or something of the sort. But the Antipodean neurologists have no need of such hypotheses.

If there is no way of explaining to the Antipodeans our problems and theories about mind and body—no way of making them see that this is the paradigm case of an ontological divide—we ought to be prepared to face up to the possibility that the "materialist" Antipodeans (as opposed to the more charitable "epiphenomenalist" ones) are right: we have just been reporting neurons when we thought we were reporting raw feels. It was just a happenstance of our cultural development that we got stuck so long with placeholders. It is as if, while perfecting many sublunary disciplines, we had never developed astronomy and had remained pre-Ptolemaic in our notions of what was above the moon. We would doubtless have many complicated things to say about holes in the black dome, movements of the dome as a whole, and the like—but once we were clued in we could redescribe what we had been reporting easily enough.

At this point, however, there is a familiar objection to be dealt with. It is expressed in such remarks as the following:

> . . . in the case of stabbing pains, it is not possible to hold that the micro-picture is the real picture, that perceptual appearances are only a coarse duplication, for in this case we are dealing with the perceptual appearances themselves, which cannot very well be a coarse duplicate of themselves.[3]

> It is all very well to claim that hurtfulness is how activity of the C-fibers in the cortex appears, that the smell of onion is how the shape of onion molecules appears to a human with a normal nasal system. . . . This deals with the pain, smell or color apprehended and, relegating it to the category of appearance, renders it ontically neutral.

[3] Richard Brandt, "Doubts about the Identity Theory," in *Dimensions of Mind*, ed. Sidney Hook (New York, 1961), p. 70.

> But it leaves us with a set of *seemings*, acts of imperfect apprehension, in which the phenomenal properties are grasped. So we must ask the new question: Is it possible that things can *seem to be* in a certain way to a merely material system? Is there a way in which acts of imperfect apprehension can be seen to be ontically neutral?
>
> . . . The materialist account of real men can find no place for the fact that our imperfect apprehension is by phenomenal property and not by, for example, beliefs just spontaneously arising.[4]

This objection common to Brandt and Campbell seems at first blush to be that one can only misdescribe things if one is not a "merely material system"—for such systems cannot have things appear to them differently from what they are. But this will not do as it stands, for, as I suggested earlier, the distinction between reality and appearance seems merely the distinction between getting things right and getting things wrong—a distinction which we have no trouble making for simple robots, servo-mechanisms, etc. To make the objection plausible we must say that "appearance" in the present context is a richer notion—one which has to be explicated by the notion of "phenomenal property." We must hold some principle like:

> (P) Whenever we make an incorrigible report on a state of ourselves, there must be a property we are presented with which induces us to make the report.

But this principle, of course, enshrines the Cartesian notion that "nothing is closer to the mind than itself," and involves an entire epistemology and metaphysics, a specifically dualistic one.[5] So it is not surprising, once we have encapsu-

[4] Keith Campbell, *Body and Mind* (New York, 1960), pp. 106-107, 109.

[5] George Pitcher has worked out an account of the linguistic behavior we display in reporting pains without using such a premise. Pitcher takes pains to be reports of damaged peripheral tissue, whereas the Antipodeans take them to be reports of states of the central nervous

lated this view in the notion of "phenomenal property," that "the materialist account . . . can find no place for the fact that our imperfect apprehension is by phenomenal property."

Still, we must ask whether there is some pre-philosophical intuition which is preserved in (P) and which can be separated from the Cartesian picture. What exactly is the difference between misdescribing something like a star and misdescribing something like a pain? Why does the former seem obviously possible and the latter unimaginable? Perhaps the answer goes something like this. We expect the star to look the same even after we realize that it is a faraway ball of flame rather than a nearby hole, but the pain ought to feel different once we realize that it is a stimulated C-fiber, for the pain *is* a feeling, as the star is *not* a visual appearance. If we give this answer, however, we are still stuck with the notion of "feeling" and with the puzzle about whether the Antipodeans have any feelings. What, we must ask, is the difference between feeling a pain and simply reacting to a stimulated C-fiber with the vocable "pain," avoidance-behavior, and the like? And here we are inclined to say: no difference at all from the outside, but all the difference in the world from the inside. The difficulty is that there will

system. In his view, it is a mistake to think of the common-sense concept of pain as the concept of a mental particular. I would want to say that it *is* the concept of a mental particular, but claim that his analysis of the epistemological status of pains applies, *mutatis mutandis*, equally well wherever one stands on this question. See Pitcher, "Pain Perception," *Philosophical Review* 79 (1970), 368-393. Pitcher's general strategy is a defense of direct realism and is also found in his *A Theory of Perception* (Princeton, 1971) and in D. M. Armstrong, *Perception and the Physical World* (London and New York, 1961) and *A Materialist Theory of the Mind* (London and New York, 1968). This strategy seems to me essentially right, and enough to show that the mental-particular view is optional. But I am dubious about Pitcher's and Armstrong's metaphilosophical stance, which would make this view a philosopher's misconstrual of what we believe, rather than a correct account of what we believe (but need not continue to believe).

never be any way in which we can explain this difference to the Antipodeans. The materialist Antipodeans think that we don't have any feelings, because they do not think there is such a thing as "feeling." The epiphenomenalist Antipodeans think that there may be such things, but cannot imagine why we make such a fuss about having them. The Terran philosophers who think that Antipodeans do have feelings but don't know it have reached the terminal stage of philosophizing mentioned by Wittgenstein: they just feel like uttering an inarticulate sound. They cannot even say to the Antipodeans that "it's different for us on the inside" because the Antipodeans do not understand the notion of "inner space"; they think "inside" means "inside the skull." *There*, they rightly remark, it *isn't* different. The Terran philosophers who think that the Antipodeans don't have feelings are in a better position only because they feel it beneath their dignity to argue with mindless beings about whether they have minds.

We seem to be getting nowhere with pursuing the objection offered by Brandt and Campbell. Let us try another tack. In the materialist view, every appearance of anything is going to be, in reality, a brain-state. So, it would seem, the materialist is going to have to say that the "coarse" duplicate of a brain-state (the way stimulated C-fibers feel) is going to be another brain-state. But, we may then say, let that other brain-state be the referent of "pain" rather than the stimulated C-fibers. Every time the materialist says "but that's just our description of a brain-state," his opponent will reply, "Okay, let's talk about the brain-state which is the 'act of imperfect apprehension' of the first brain-state."[6] And so the materialist seems to be pressed ever backward—with the mental cropping up again wherever error does. It is as if man's Glassy Essence, the Mirror of Nature, only became visible to itself when slightly clouded. A neural

[6] I owe this way of putting the Brandt-Campbell point to Thomas Nagel.

system can't have clouds but a mind can. So minds, we conclude, cannot be neural systems.

Consider now how the Antipodeans would view "acts of imperfect apprehension." They would see them not as cloudy portions of the Mirror of Nature but as a result of learning a second-rate language. The whole notion of incorrigibly knowable entities, as opposed to being incorrigible about how entities seem to be—the notion of "seemings" as themselves a kind of entity—strikes them as a deplorable way of speaking. The whole Terran vocabulary of "acts of apprehension," "cognitive states," "feelings," etc. strikes them as an unfortunate turn for a language to have taken. They see no way of getting us out of it except by proposing that we raise some of our children to speak Antipodean and see whether they don't do as well as a control group. The Antipodean materialists, in other words, see our notion of "mind and matter" as a reflection of an unfortunate linguistic development. The Antipodean epiphenomenalists are baffled by the question "What is the neural input to the Terrestrial speech center which produces pain reports as well as C-fiber reports?" Those Terrestrial philosophers who think that Antipodeans do have feelings think that the Antipodean language is "inadequate to reality." Those Terrestrial philosophers who think that the Antipodeans don't have feelings rest their case on a theory of language-development according to which the first things named are the things "better known to us"—raw feels—so that the absence of a name for feeling entails the absence of feeling.

To sharpen the issue a bit further, perhaps we may drop from consideration the Antipodean epiphenomenalists and the Terrestrial skeptics. The former's problem about the neurology of pain reports seems insoluble; if they are to continue charitably to ascribe states to Earthmen which are unknown to Antipodeans they will have to swallow a whole dualistic system, irrefutable by further empirical inquiry, in order to explain our linguistic behavior. As for the Terran skeptic's claim that the Antipodeans have no raw

feels, this is based entirely on the a priori dictum that one cannot have a raw feel and lack a word for it. Neither intellectual position—the extreme charity of the Antipodean epiphenomenalist and the parochial distrust of the Terrestrial skeptic—is attractive. We are left with the Antipodean materialist saying "They think they have feelings but they don't" on the one hand and Terrestrial philosophers saying "They have feelings but don't know it" on the other hand. Is there a way out of this impasse, given that every empirical result (brain-switching, etc.) seems to weigh equally on both sides? Are there powerful philosophical methods which will cut through the problem and either settle it or offer some happy compromise?

3. Incorrigibility and Raw Feels

One philosophical method which will do no good at all is "analysis of meanings." Everybody understands everybody else's meanings very well indeed. The problem is that one side thinks there are too many meanings around and the other side too few. In this respect the closest analogy one can find is the conflict between inspired theists and uninspired atheists. An inspired theist, let us say, is one who "just knows" that there are supernatural beings which play certain explanatory roles in accounting for natural phenomena. (They are not to be confused with natural theologians—who offer the supernatural as the best explanation of these phenomena.) Inspired theists have inherited their picture of the universe as divided into two great ontological realms—the supernatural and the natural—along with their language. The way they talk about things is inextricably tied up with—or at least strikes them as inextricably tied up with—references to the divine. The notion of the supernatural does not strike them as a "theory" any more than the notion of the mental strikes us as a "theory." When they encounter atheists they view them as people who don't know what's going on, although they admit that atheists

seem able to predict and control natural phenomena very nicely indeed. ("Thank heaven," they say, "that we are not as those natural theologians are, or we too might lose touch with the real.") The atheists view these theists as having too many words in their language and too many meanings to bother about. Enthusiastic atheists explain to inspired theists that " all there *really* is is . . . ," and the theists reply that one should realize that there are more things in heaven and earth. . . . And so it goes. The philosophers on both sides may analyze meanings until they are blue in the face, but all such analyses are either "directional" and "reductive" (e.g., "noncognitive" analyses of religious discourse, which are the analogue of "expressive" theories of pain reports) or else simply describe alternative "forms of life," culminating in nothing more helpful than the announcement: "This language-game is played." The theists' game is essential to their self-image, just as the image of man's Glassy Essence is essential to the Western intellectual's, but neither has a larger context available in which to evaluate this image. Where, after all, would such a context come from?

Well, perhaps from philosophy. When experiment and "meaning analysis" fail, philosophers have traditionally turned to system-building—inventing a new context on the spot, so to speak. The usual strategy is to find a compromise which will enable both those who favor Occam's Razor (e.g., materialists, atheists) and those who cling to what they "just know" to be viewed indulgently as having achieved "alternate perspectives" on some larger reality which philosophy has just adumbrated. Thus some tender-minded philosophers have risen above the "warfare between science and theology" and seen Bonaventure and Bohr as possessing different, noncompetitive "forms of consciousness." The question "consciousness of *what?*" is answered by something like "the world" or "the thing-in-itself" or "the sensible manifold" or "stimulations." It does not matter which of these is offered, since all are terms of art de-

signed to name entities with no interesting features save placid neutrality. The analogue of this tactic among tough-minded philosophers of mind is neutral monism, in which the mental and the physical are seen as two "aspects" of some underlying reality which need not be described further. Sometimes we are told that this reality is intuited (Bergson) or is identical with the raw material of sensation (Russell, Ayer), but sometimes it is simply postulated as the only means of avoiding epistemological skepticism (James, Dewey). In no case are we told anything about it save that "we just know what it's like" or that reason (i.e., the need to avoid philosophical dilemmas) requires it. Neutral monists like to suggest that philosophy has discovered, or should look for, an underlying substrate, in the same way in which the scientist has discovered molecules beneath elements, atoms beneath molecules, and so on. But in fact the "neutral stuff" which is neither mental nor physical is not found to have powers or properties of its own, but simply postulated and then forgotten about (or, what comes to the same thing, assigned the role of ineffable datum).[7] This tactic cannot help in coping with the question which the tough-minded Terrestrial philosophers raised

[7] Urging that philosophers need to do more than this, Cornelius Kampe has suggested that the mind-body identity theory will make sense only if we provide "a theoretical framework (or an ontology for the common idiom) of such a type as to provide a link for the two diverse phenomena whose identity is asserted." His motive for this revival of neutral monism is his belief that making sense of an identity theory requires that "the subjective-objective distinction must be abandoned, as must the privileged status of first-person introspective reports." Such a change, he says, would "drastically affect the logic of our language." I think Kampe is right that giving up the subjective-objective distinction would have such a drastic effect, but wrong in thinking that giving up privileged access would. As I think Sellars has shown, and as I have been arguing here, the subjective-objective distinction (the notion of "seems") can get along quite well without the notions of "mind," "phenomenal property," etc. (Cf. Kampe, "Mind-Body Identity: A Question of Intelligibility," *Philosophical Studies* 25 [1974], 63-67.)

about the Antipodeans: Do they have raw feels or don't they?

The problem about the Antipodeans can be summarized as follows:

1. It is essential to raw feels that they be incorrigibly knowable

together with

2. There is nothing which the Antipodeans think themselves incorrigible about

seems to leave us either with

3. The Antipodeans do not have raw feels

or with

4. The Antipodeans do not know about their own incorrigible knowledge.

The trouble with (3) is that the Antipodeans have pretty much the behavior, physiology, and culture that we do. Further, we can train Antipodean infants to report raw feels, and take themselves to be incorrigible about them. These considerations seem to drive us toward (4). But (4) sounds silly, and needs at least to be softened to

4′. The Antipodeans do not know about their own capacity for incorrigible knowledge

which is a little odd but at least has a few parallels. (Compare "John XXIII had to be convinced by argument of his own infallibility upon succeeding to the papacy.") However, if we press (4′), the teachability of Antipodean infants seems to leave us up in the air between

5. The Antipodeans can be taught to recognize their own raw feels

and

5′. The Antipodeans can be taught, thanks to the presence of neural concomitants of raw feels, to simulate reports of raw feels without actually having any.

One might hope to resolve this new dilemma by finding a bilingual Antipodean. But the bilingual does not have "inside" knowledge about the meanings of the foreign expressions; he just has the same sort of theory which the lexicon-maker has. Consider an adult Antipodean who has come to speak English. He says, "I am in pain" or the Antipodean for "My C-fibers are firing," depending on which he's speaking. If a Terran interlocutor tells him that he really isn't in pain he points out that the remark is a deviant utterance, and claims privileged access. When Antipodean interlocutors show him that his C-fibers aren't in fact firing he says something like "That's funny; they certainly seem to be. That's why I told the Terrans I was in pain," or perhaps something like "That's funny; I'm certainly in what the Terrans call 'pain,' and that never happens except when my C-fibers are firing." It is hard to see that he would have any strong preference for either locution, and harder to see that philosophers could make anything out of a preference if he had one. Once again, we seem driven to the rhetorical question "But what does it *feel* like?"—to which the bilingual Antipodean replies, "It feels like pain." When asked, "Doesn't it also feel like C-fibers?" he explains that there is no concept of "feeling" in Antipodean, and so it would not occur to him to say that he *felt* his C-fibers firing, although of course he is aware of it whenever they do.

If this seems paradoxical, it is presumably because we think that "noninferential awareness" and "feeling" are pretty well synonymous. But pointing this out is no help. If we treat them as synonymous, then of course Antipodean *does* have the concept of a state called "feeling," but it still doesn't have the concept of "feelings" as intentional objects of knowledge. Antipodean has the verb but not the noun, so to speak. An accommodating Antipodean can note that his

language can express the notion of "state such that one cannot be mistaken in thinking one is in it"—namely, the state of it seeming to one that . . .—but still be puzzled about whether these states are the same things as the pains and other raw feels in which the Terrans are so interested. On the one hand, it seems all they *could* be talking about, for he remembers having learned to say "pain" when and only when his C-fibers seem to be firing. On the other hand, the Terrans insist that there is a difference between being in a state such that it seems to one that one is . . . and having a raw feel. The former state is an epistemic position toward something about which doubt is possible. The latter state automatically puts one in an epistemic position toward something about which doubt is impossible.

So the dilemma seems to boil down to this: We must affirm or deny

6. any report of how something seems to one is a report of a raw feel.

The only ground for affirming it seems to be that it is a corollary of the converse of (1), that is:

7. It is essential to whatever is incorrigibly knowable that it be a raw feel.

But (7) is just a form of the principle invoked by the Brandt-Campbell objection above, viz.:

(P) Whenever we make an incorrigible report on a state of ourselves, there must be a property which we are presented with which induces us to make the report

and in this principle everything turns on the notion "presented with"—a notion which harks straight back to the metaphors of the "Eye of the Mind," "presence to consciousness," and the like, which are in turn derived from the initial image of the Mirror of Nature—of knowledge as a set of immaterial representations. If we adopt this prin-

ciple, then, oddly enough, we can no longer be skeptics: the Antipodeans automatically have raw feels. We *must* choose (5) over (5′). Since we do not contest that it seems to some Antipodean that his stomach is cramped or his C-fibers firing, and since we grant incorrigibility to such reports, we must grant that he has some raw feels which are the "basis" for his seems-statements and which he could be trained to report by learning an appropriate vocabulary. But this means, paradoxically enough, that a species of behaviorism is entailed by the very principle that incarnates the Cartesian image of the Eye of the Mind—the very image which has often been accused of leading to the "veil of ideas" and to solipsism. We should only be able to be skeptics and assert (5′)—that simulation might be all the Antipodeans could do—by holding that when the Antipodeans made seems-statements they were not really meaning what we meant by them, and that the deviance, in Antipodean, of the expression "You may be mistaken in saying that it seems to you that your C-fibers are firing" does not suffice to show that the Antipodeans have any incorrigible knowledge. That is, we should have to reconstrue the behavior which we initially took to be exhibited, and base our skepticism about their raw feels on a more general skepticism about their possession of knowledge (or of some kinds of knowledge). But it is difficult to see how we could make skepticism about this plausible except on some antecedent conviction that they were mindless—a conviction which would a fortiori rule out raw feels. So skepticism here will have to be groundless and Pyrrhonian.[8] On the other

[8] I want to distinguish between "mere," or Pyrrhonian, skepticism and the specifically "Cartesian" form of skepticism which invokes the "veil of ideas" as a justification for a skeptical attitude. "Pyrrhonian" skepticism, as I shall use the term, merely says, "We can never be certain; so how can we ever know?" "Veil of ideas" skepticism, on the other hand, has something more specific to say, viz., "Given that we shall never have certainty about anything except the contents of our own minds, how can we ever justify an inference to a belief about anything else?" For a discussion of the intertwining of these two forms of skepticism, see Richard Popkin, *The History of Scepticism from Erasmus to Descartes* (New York, 1964).

hand, if we deny (6)—if we disengage seeming from the having of mental states and abandon the Cartesian pictures—then we have to face up to the possibility that we ourselves never had any feelings, any mental states, any minds, any Glassy Essence. This paradox seems so overwhelming as to drive us right back to (P) and the Mirror of Nature.

So the problem comes down to a choice among three troubling possibilities. We have to either share our Glassy Essence with any being which seems to speak a language containing seems-statements, or become Pyrrhonian skeptics, or else face up to the possibility that this essence was never ours. If we grant (7) above—the premise which makes being a raw feel essential to being an object of incorrigible knowledge—then we must admit either (a) that the Antipodean language, just by virtue of containing some incorrigible reports, is about raw feels, or (b) that we shall never know whether the Antipodeans speak a language just because we shall never know whether they have raw feels, or (c) that the whole issue about raw feels is a fake because the example of the Antipodeans shows that we never had any raw feels ourselves.

These three possibilities correspond roughly to three standard positions in the philosophy of mind—behaviorism, skepticism about other minds, and materialism. Rather than adopt any of these three, however, I suggest that we deny (7), and with it (P). That is, I suggest that we abandon the notion that we possess incorrigible knowledge by virtue of a special relation to a special kind of object called "mental objects." This suggestion is a corollary of Sellars's attack on the Myth of the Given. I shall present that attack in more detail in chapter four, but here I merely note that this myth is the notion that such epistemic relations as "direct knowledge" or "incorrigible knowledge" or "certain knowledge" are to be understood on a causal, para-mechan-

ical model, as a special relation between certain objects and the human mind which enables knowledge to take place more easily or naturally or quickly. If we think of incorrigible knowledge simply as a matter of social practice —of the absence of a normal rejoinder in normal conversation to a certain knowledge-claim—then no principle like (7) or (P) will seem plausible.

In the last two sections I have been treating "mental object" as if it were synonymous with "incorrigibly knowable object," and thus as if to have a mind were the same thing as having incorrigible knowledge. I have disregarded immateriality and the ability to abstract, which were discussed in chapter one, and intentionality, which will be discussed in chapter four. My excuse for pretending that the mind is nothing but a set of incorrigibly introspectible raw feels, and that its essence is this special epistemic status, is that the same pretense is current throughout the area called "philosophy of mind." This area of philosophy has come into existence in the thirty years since Ryle's *The Concept of Mind*. The effect of that book was to make issues about minds and bodies turn almost entirely on the cases which resisted Ryle's own logical behaviorist attempt to dissolve Cartesian dualism—namely, raw feels. Wittgenstein's discussion of sensations in *Philosophical Investigations* seemed to offer the same sort of attempt at dissolution. Thus many philosophers have taken it for granted that "the mind-body problem" was the question of whether raw feels could be viewed as dispositions to behave. Thus the only possibilities have seemed to be the ones I have just cited: (a) granting that Ryle and Wittgenstein were right, and that there are no mental objects, (b) saying that they were wrong, and that therefore Cartesian dualism stands intact, with skepticism about other minds a natural consequence, and (c) some form of mind-brain identity theory, according to which Ryle and Wittgenstein were wrong, but Descartes is not thereby vindicated.

The effect of setting up the issues in this way is to focus

on pains, while paying less attention to the side of the mind which is, or should be, of more concern to epistemology—beliefs and intentions. (The balance has been somewhat redressed in recent years thanks to philosophers of mind who try to build bridges with empirical psychology. Their work will be discussed in chapter five.) But it is still the case that "the mind-body problem" is thought of primarily as a problem about pains, and the distinctive point about pains is just the one mentioned by Kripke—that there seems no such thing as an appearance-reality distinction in regard to our knowledge of them. In fact, as I have tried to show in chapter one, this is only one of several "mind-body problems," each of which has contributed to the fuzzy notion that there is something especially mysterious about man which makes him capable of knowing, or of certain special sorts of knowing.

For the remainder of this chapter, however, I shall try to support my claim that we should drop (P) and thus be neither dualists, skeptics, behaviorists, nor "identity-theorists." I do not know how to argue against (P) directly, since the claim that incorrigible knowledge is a matter of being presented with a phenomenal property is not so much a claim as an abbreviation for an entire theory—a whole set of terms and assumptions which center around the image of mind as mirroring nature, and which conspire to give sense to the Cartesian claim that the mind is naturally "given" to itself. It is this image itself which has to be set aside if we are to see through the seventeenth-century notion that we can understand and improve our knowing by understanding the workings of our mind. I hope to show the difference between setting it aside and adopting any of the positions which presuppose this image. So the remainder of this chapter is devoted to behaviorism, skepticism, and the mind-body identity theory, in an attempt to differentiate my position from each of these. In the concluding section of the chapter—"Materialism without Identity"—I attempt to say something more positive, but this attempt needs to

be linked up with the discussion of other "mind-body problems" in chapter one in order to appear plausible.

4. Behaviorism

Behaviorism is the doctrine that talk of "inner states" is simply an abbreviated, and perhaps misleading, way of talking of dispositions to behave in certain ways. In its Rylean or "logical" form—with which I shall be concerned in what follows—its central doctrine is that there is a necessary connection between the truth of a report of a certain raw feel and a disposition to such-and-such behavior. One motive for holding this view is a distrust of what Ryle called "ghosts in machines," the Cartesian picture of people, and another is the desire to prevent the skeptic about other minds from raising the question of whether the person writhing on the floor has feels of the sort which the skeptic himself would have when he writhes. In the logical behaviorist view, reports of such feels are to be taken not to refer to nonphysical entities, and perhaps not to any entities at all save to the writhing or the disposition to writhe.

This doctrine has been attacked on the ground that there seems no way to fill in a description of the requisite disposition to behave without giving infinitely long lists of possible movements and noises. It has also been attacked on the ground that whatever "necessity" there is in the area is not a matter of "meaning" but simply an expression of the fact that we customarily explain certain behavior by reference to certain inner states—so that the necessity is no more "linguistic" or "conceptual" than that which connects the redness of the stove to the fire within. Finally, it has been attacked as the sort of philosophical paradox which would only occur to a mind obsessed with instrumentalist or verificationist dogma—eager to reduce all unobservables to observables in order to avoid any risk of believing in something unreal.

All these criticisms are, I think, quite justified. The classic

statements of logical behaviorism do indeed presuppose just the distinctions between observation and theory and between language and fact which philosophers would, as I shall argue in chapter four, do well to give up. But the feeling that the behaviorist is on to something remains. One point which he has going for him is that it seems absurd to suggest that we might someday, after years of fruitful conversation with the Antipodeans, have ground for saying, "Ah, no raw feels; so no minds; so no language and not persons after all." The suggestion that we might find ourselves compelled to say that they had no raw feels makes us ask whether we can even imagine what such a compulsion could be like. It also makes us realize that even if we somehow were so compelled, we should almost certainly not draw the suggested inferences. On the contrary, we might begin to share the Antipodeans' bewilderment about why we had cared so much about this question. We should begin to appreciate the quizzical attitude which the Antipodeans adopt toward the whole topic—the same attitude with which the Polynesians viewed missionaries' preoccupation with the question "Are these descendants of Shem or of Ham?" The behaviorist's strong point is that the more one tries to answer them the more pointless the tough-minded philosopher's questions "Minds or no minds?" "Raw feels or no raw feels?" seem to become.

But this good point begins to go bad as soon as it is put as a thesis about "necessary connections" established by "analysis of meanings." Ryle's insight was frustrated by the positivistic epistemology he inherited. Instead of saying that incorrigible knowledge was just a matter of what practices of justification were adopted by one's peers (the position which I shall call "epistemological behaviorism" in chapter four), he was led to say that a certain type of behavior formed a necessary and sufficient condition for the ascription of raw feels, and that this was a fact about "our language." He then was confronted by a stubborn problem. The fact that our language licensed the inference to the