Chapter 3—
The Semantical Problem

Where do the terms of our common-sense psychological vocabulary get their meanings? This apparently innocent question is important for at least three reasons. Psychological terms form a crucial test case for theories of meaning in general. The semantical problem is closely bound up with the ontological problem, as we saw in the first chapter. And it is even more closely bound up with the epistemological problem, as we shall see in the next chapter.

In this chapter, we shall explore the cases for and against each of the three main theories currently at issue. The first says that the meaning of any common-sense psychological term (of most of them, anyway) derives from an act of *inner ostension*. A second insists that their meaning derives from *operational definitions*. And a third claims that the meaning of any such term derives from its place in a *network of laws* that Constitute 'folk' psychology. Without further ado, let us address the first theory.

1—
Definition by Inner Ostension

One way to introduce a term to someone's vocabulary—"horse", or "fire engine", for example—is just to show the person an item of the relevant type, and say something like, "*That* is a horse," or "*This* is a fire engine." These are instances of what is called *ostensive definition*. One expects the hearer to notice the relevant features in the situation presented, and to be able to reapply the term when a new situation also contains them.

Of course, both of the expressions cited could have been introduced in another way. One could just have said to the hearer, "A horse is a large, hoofed animal used for riding." Here one gives the meaning of the term by connecting it in specific ways with other terms in the hearer's vocabulary. Such term introductions range from the explicit and complete ("An isosceles triangle is a three-sided closed plane figure with at least two equal sides") to the partial and casual ("Energy is what makes our cars run and keeps our lights burning"). But not all

terms get their meaning in this way, it is often said. Some terms can get their meaning only in the first way, by direct ostension. Terms like "red", "sweet", and "warm", for example. Their meaning is not a matter of the relations they bear to other terms; it is a matter of their being directly associated with a specific quality displayed by material objects. Thus speaks orthodox semantic theory and common sense alike.

What of the terms in our common-sense psychological vocabulary? When one thinks of terms like "pain", "itch", and "sensation of red", ostension seems the obvious source of meaning. How could one possibly know the meaning of any of these terms unless one had actually had a pain, or an itch, or a sensation of red? Prima facie, it seems one could not. Call this "the standard view".

While the standard view may be correct for a significant class of psychological terms, it is clearly not correct for all such terms, nor even for the majority. Many important types of mental states have no qualitative character at all, or none that is relevant to their type-identity. Consider the variety of different beliefs, for example: the belief that *P*, the belief that *Q*, the belief that *R*, and so on. We have here a potential infinity of importantly different states. One could not possibly master the meaning of each expression by learning, one by one, a qualitative character peculiar to each state. Nor does each have a distinct quale anyhow. And the same goes for the potential infinity of distinct thoughts that *P*, and desires that *P*, and fears that *P*, and for all of the other 'propositional attitudes' as well. These are perhaps the most central expressions in our common-sense framework, and they are distinguished by a role-playing element, the sentence *P*, not by some introspectible quale (= 'phenomenological quality'). Their meaning must derive from some other source.

Clearly the standard view cannot be the whole story about the meaning of psychological predicates. Further, the standard view is suspect even in its most plausible cases. Among those mental states that are associated with qualia, not all types have a *uniform* quale. In fact, very few do, if any. Consider the term "pain", and think of the wide variety of substantially different sensations included under that term (think of a headache, a burn, a piercing noise, a blow to the kneecap, and so on). Granted, all of these qualia are similar in causing a reaction of dislike in the victim, but this is a *causal/relational* property common to all pains, not a shared quale. Even sensations-of-red show a wide variation through many shades and hues, bordering on brown, orange, pink, purple, or black at their several extremes. Granted, intrinsic similarities do something to unify this diffuse class, but it seems clear that the class of sensations-of-red is equally delimited by the fact that

sensations-of-red typically result from viewing such standard examples as lips, strawberries, apples, and fire engines. That is, they are united by their shared causal/relational features. The idea of meaning being exhausted by a single, unambiguous quale seems to be a myth.

Are we certain that knowing the quale is even necessary to knowing the meaning? It has been argued that someone who has never been in pain (perhaps because of some fault in his nervous system) could still know the meaning of the word "pain" and use it conversation, explanation, and prediction, just as we use it in describing others. Granted, he would not know what pain *feels* like, but he could still know all of its causal/relational properties, and hence would know as well as we do what kind of state pain is. There would remain *something* he did not know, but it is not clear that the something is the meaning of the word "pain".

If the meaning of terms like "pain" and "sensation-of-red" really were exhausted by their association with an inner quale, then we would :be hard pressed to avoid a *semantic solipsism*. (Solipsism is the thesis that all knowledge is impossible save for knowledge of one's immediate self.) Since each one of us can experience only one's *own* states of consciousness, it would then be impossible for anyone to tell whether or not one's own meaning for "pain" is the same as anyone else's. And surely it is an odd theory of meaning that entails that no one ever understands what anyone else means.

These doubts about the standard 'inner ostension' theory of meaning have prompted philosophers to explore other approaches. The first serious attempt to articulate and to defend an alternative theory was provided by the philosophical behaviorists, whom we met in the preceding chapter. These thinkers advanced a further argument against the standard view, which we shall now examine.

2—
Philosophical Behaviorism

According to the behaviorists, the meaning of any mental term is fixed by the many relations it bears to certain other terms: terms for publicly observable circumstances and behaviors. In its clearest formulations, behaviorism pointed to purely dispositional terms like "soluble" and "brittle" as semantic analogues for mental terms, and it pointed to operational definitions as the structures whereby the meanings of mental terms could be made explicit. The details of this view were outlined in chapter 2.2, so I shall not repeat them here.

A major problem for behaviorism was the insignificant role it assigned to the qualia of our mental states. But we have just seen some good reasons for reestimating (downward) the importance standardly assigned to qualia. And one of the most influential philosophers in the behaviorist tradition, Ludwig Wittgenstein, had a further argument against the standard view: the *private language argument*.

Despite the consequence of semantic solipsism, many defenders of the standard view were prepared to live with the idea that one's vocabulary for sensations was an inescapably *private* language. Wittgenstein attempted to show that a necessarily private language was completely impossible. The argument ran as follows. Suppose you attempt to give meaning to a term "W" solely by associating it with a certain sensation you feel at the time. At a later time, upon feeling a sensation, you may say, "There is another W." But how can you determine whether you have used the term correctly on this occasion? Perhaps you misremember the first sensation, or carelessly see a close similarity between the second and first where in fact there is only a faint and distant resemblance. If the term "W" enjoys no meaning connections whatsoever with *other* phenomena, such as certain standard causes and/or effects of the kind of sensation at issue, then there will be absolutely no way to distinguish between a correct use of "W" and an incorrect use of "W". But a term whose proper application is forever beyond determination is a meaningless term. A necessarily private language is therefore impossible.

This argument gave behaviorists much encouragement in their attempts to define our common expressions for mental states in terms of their connections with publicly observable circumstances and behaviors. Despite the encouragement, those attempts never really succeeded (as we saw in chapter 2.2), and frustration gathered quickly. Perhaps this should have been expected, because Wittgenstein's private language argument draws a stronger conclusion than its premises justify. If a check on correct application is what is required for meaningfulness, then all that one's understanding of "W" need include is some con-

nections between the occurrence of the W-sensation and the occurrence of *other* phenomena. Those other phenomena *need* not be publicly observable phenomena: they can be other mental states, for example, and still serve as checks on the correct application of "W".

What Wittgenstein's argument should have concluded, therefore, is just that no term can be meaningful in the absence of systematic connections with other terms. Meaning, it appears, is something a term can enjoy only in the context of a network of other terms, terms connected to one another by means of general statements that contain them. If Wittgenstein and the behaviorists had drawn this slightly weaker conclusion, then perhaps philosophers might have arrived at the semantic theory of the following section more swiftly than they did.

Suggested Readings

Malcolm, Norman, "Wittgenstein's *Philosophical Investigations*," the *Philosophical Review*, vol. LXIII (1954). 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 and M. Scriven (Minneapolis: University of Minnesota Press, 1958). Reprinted in *The Philosophy of Mind*, ed. V. C. Chappell (Englewood Cliffs, NJ: Prentice-Hall, 1962).

Hesse, Mary, "Is There an Independent Observation Language?" in *The Nature and Function of Scientific Theories*, ed. R. Colodny (Pittsburgh: Pittsburgh University Press, 1970). See especially pp. 44–45.

3—
The Theoretical Network Thesis and Folk Psychology

The view to be explored in this section can be stated as follows. Our common-sense terms for mental states are the *theoretical terms* of a theoretical framework (folk psychology) embedded in our commonsense understanding, and the meanings of those terms are fixed in the same way as are the meanings of theoretical terms in general. Specifically, their meaning is fixed by the set of laws/principles/generalizations in which they figure. In order to explain this view, let me back up a few steps and talk about theories for a few moments.

The Semantics of Theoretical Terms

Consider large-scale theories, such as those found in the physical sciences: chemical theory, electromagnetic theory, atomic theory, thermodynamics, and so on. Typically, such a theory consists of a set of sentences—usually general sentences or *laws*. These laws express the relations that hold between the various properties/values/classes/entities whose existence is postulated by the theory. Such properties and entities are expressed or denoted by the set of *theoretical terms* peculiar to the theory in question.

Electromagnetic theory, for example, postulates the existence of electric charges, electric force fields, and magnetic force fields; and the laws of electromagnetic theory state how these things are related to one another and to various observable phenomena. To fully understand the expression "electric field" is to be familiar with the network of theoretical principles in which that expression appears. Collectively, they tell us what an electric field is and what it does.

This case is typical. Theoretical terms do not, in general, get their meanings from single, explicit definitions stating conditions necessary and sufficient for their application. They are implicitly defined by the network of principles that embed them. Such casual 'definitions' as one does find given (for example, "The *electron* is the unit of electricity") usually give only a small part of the term's significance, and are always subject to falsification in any case (for example, it now appears that the *quark* may be the unit of electricity, with a charge one-third that of the electron). Call this view the *network theory of meaning*.

The Deductive-Nomological Model of Explanation

The laws of any theory do more, however, than just give sense to the theoretical terms they contain. They also serve a predictive and an *explanatory* function, and this is their main value. Which raises the question: What is it to give an explanation of an event or state of affairs, and how do theories make this possible?

We may introduce the conventional wisdom on this point with the following story.

In my laboratory there is an apparatus consisting of a long metal bar with two facing mirrors, one attached to each end. The point of the bar is to keep the mirrors a precise distance apart. One morning, while remeasuring the distance just prior to performing some experiment, my assistant notices that the bar is now longer than it was, by about one millimeter.

"Hey," he announces, "this bar has expanded. Why is that?"

"Because I heated it," I explain.

"Y-e-s?" he queries, "what's that got to do with anything?"

"Well, the bar is made of copper," I explain further.

"Y-e-s?" he persists, "and what's that got to do with it?"

"Well, all copper expands when heated," I reply, suppressing exasperation.

"A-h-h, I see," he says, as the light finally dawns.

If, after my final remark, my assistant had still failed to understand, then I should have to fire him, because the explanation of why the bar expanded is now complete, and even a child should get it. We can see why, and in what sense, it is complete by looking at the assembled information my explanation contained.

1. All copper expands when heated.

2. This bar is copper.

3. This bar is heated.

———

4. This bar is expanded.

The reader will notice that, collectively, the first three propositions *deductively entail* the fourth proposition, the statement of the event or state of affairs to be explained. The bar's expansion is an inevitable consequence of the conditions described in the first three propositions.

We are looking here at a valid deductive *argument*. An explanation, it seems, has the form of an argument, an argument whose premises (the *explanans*) contain the explanatory information, and whose conclusion (the *explanandum*) describes the fact to be explained. Most important, the premises include a *nomological* statement—a law of nature, a general statement expressing the patterns to which nature adheres. The other premises express what are commonly called "initial conditions", and it is these that connect the law to the specific fact in need of explanation. In sum, to explain an event or state of affairs is to deduce its description from a law a nature. (Hence the name, "the

deductive-nomological model of explanation".) The connection between comprehensive theories and explanatory power is now easy to see.

The *prediction* of events and states of affairs, we should note, follows essentially the same pattern. The difference is that the conclusions of the relevant arguments are in the future tense, rather than in the past or present tense. Notice also a further point. When voicing an explanation in ordinary life, one hardly ever states every premise of the relevant argument. (See my first response to my assistant.) There is generally no point, since one can assume that one's hearers already possess most of the relevant information. What one gives them is just the specific piece of information one presumes they are missing (for example, "I heated it"). Most explanations, as voiced, are only explanation sketches. The hearer is left to fill in what is left unsaid. Last, it should be pointed out that the 'laws' that lie behind our commonsense explanations are usually on the rough-and-ready side, expressing only a rough approximation to, or an incomplete grasp of, the true regularities involved. This is thus one further dimension in which our explanations are generally explanation sketches.

Folk Psychology

Consider now the considerable capacity that normal humans have for explaining and predicting the behavior of their fellow humans. We can even explain and predict the psychological states of other humans. We explain their behavior in terms of their beliefs and desires, and their pains, hopes, and fears. We explain their sadness in terms of their disappointment, their intentions in terms of their desires, and their beliefs in terms of their perceptions and inferences. How is it we are able to do all this?

If the account of explanation in the preceding section is correct, then each of us must possess a knowledge or a command of a rather substantial set of laws or general statements connecting the various mental states with (1) other mental states, with (2) external circumstances, and with (3) overt behaviors. Do we?

We can find out by pressing some common-sense explanations, as the explanation was pressed in the sample conversation earlier, to see what other elements are commonly left unsaid. When we do, argue the proponents of this view, we uncover literally hundreds and hundreds of common-sense generalizations concerning mental states, such as the following:

Persons tend to feel pain at points of recent bodily damage.

Persons denied fluids for some time tend to feel thirst.

Persons in pain tend to want to relieve that pain.

Persons who feel thirst tend to desire drinkable fluids.

Persons who are angry tend to be impatient.

Persons who feel a sudden sharp pain tend to wince.

Persons who are angry tend to frown.

Persons who want that *P*, and believe that *Q* would be sufficient to bring about *P*, and have no conflicting wants or preferred strategies, will try to bring it about that *Q*.

These familiar platitudes, and hundreds of others like them in which other mental terms are embedded, are what constitute our understanding of how we work. These rough-and-ready general statements or *laws* support explanations and predictions in the normal fashion. Collectively, they constitute a *theory*, a theory that postulates a range of internal states whose causal relations are described by the theory's laws. All of us learn that framework (at mother's knee, as we learn our language), and in so doing we acquire the common-sense conception of what conscious intelligence *is*. We may call that theoretical framework "folk psychology". It embodies the accumulated wisdom of thousands of generations' attempts to understand how we humans work.

To illustrate, briefly, the role that such laws play in ordinary explanations, consider the following exchange.

"Why did Michael wince slightly when he first sat down to the meeting?"

"Because he felt a sudden sharp pain."

"I see. And why did he feel a pain?"

"Because he sat on the tack I placed on his chair."

Here we have two explanations, one on the heels of the other. If each is pressed, in the manner of our initial example, the sixth and first laws on the preceding list will emerge from the presumptive background, and two deductive arguments will become apparent, showing the same pattern as the explanation of the expanded bar.

If folk psychology is literally a theory—albeit a very old theory, deeply entrenched in human language and culture—then the meanings of our psychological terms should indeed be fixed as the thesis of this section says they are: by the set of folk-psychological laws in which they figure. This view has a certain straightforward plausibility; after all, who will say that someone understands the meaning of the term "pain" if he has no idea that pain is caused by bodily damage, that people hate it, or that it causes distress, wincing, moaning, and avoidance behavior?

Qualia Again

But what of the qualia of our various psychological states? Can we really believe, as the network theory seems to require, that qualia play *no* role in the meanings of our psychological terms? The intuition that they do is extremely strong. There are at least two ways in which a defender of the network theory might try to handle this perennial intuition.

The first is just to admit that qualia do play *some* role in the meaning of *some* terms, though only a minor or secondary role at best. This concession would go a long way toward soothing our intuitions, and it is tempting just to adopt it and to declare the issue closed. But it does leave certain problems unsolved. Since the qualia of your sensations are apparent only to you, and mine only to me, *part* of the meaning of our sensation-terms will remain private, and it will still be a stubbornly open question whether any of us means the same thing by those terms.

The second compromise concedes to qualia a significant role in the introspective *application* of sensation, terms, but still attempts to deny that their role enjoys any *semantic* significance. The idea is that your introspective discrimination of a pain from a tickle, or a sensation-of-red from a sensation-of-green, is of course keyed to the qualitative character, in you, of the relevant states. Each of us learns to exploit such qualia as our states display, in order to make spontaneous observation judgments as to which states we are in. But what is strictly meant by "pain", for example, does not include any commitment to any specific qualia. The qualitative character of pains varies substantially even within a given individual; it may well vary even more widely across different individuals; and almost certainly it varies substantially across distinct biological species. Qualia, therefore, have an epistemological significance, but they are without semantic significance for terms in an intersubjective language.

Thus two competing addenda to the network theory of meaning. Which one should be adopted I leave to the reader to decide. In either case, the background lesson appears plain: the dominant, and perhaps the only, source of meaning for psychological terms is the commonsense theoretical network in which they are embedded. As with theoretical terms generally, one comes to understand them only as one learns to use the predictive and explanatory generalizations in which they figure.

General Significance

The significance of this network theory of meaning—for the mindbody problem—is as follows. The network theory is strictly consistent with all three of the current materialist positions, and it is also consistent

with dualism. It does not by itself entail or rule out any of these positions. What it does do is tell us something about the nature of the conflict between them all, and about the way in which the conflict will be resolved. The lesson is as follows.

If our common-sense framework for psychological states is literally a *theory*, then the question of the relation of mental states to brain states becomes a question of how an old theory (folk psychology) is going to be related to a new theory (matured neuroscience) which threatens in some way to displace it. The four major positions on the mind-body issue emerge as four different anticipations of how that theoretical conflict is going to be resolved. The identity theorist expects that the old theory will be smoothly reduced by a matured neuroscience. The dualist maintains that the old theory will not be reduced by a matured neuroscience, on the grounds that human behavior has nonphysical sources. The functionalist also expects that the old theory will not be reduced, but because (ironically) too many different kinds of physical systems can produce the exact causal organization specified by the old theory. And the eliminative materialist also expects that the old theory will fail to reduce, on the yet different grounds that it is simply too confused and inaccurate to win survival through intertheoretic reduction.

What is at issue here is the fate of a theory, the fate of a speculative explanatory framework, namely, our own beloved folk psychology. And it is apparent that the issue between these four possible fates is basically an empirical issue, to be settled decisively only by continuing research in the neurosciences, cognitive psychology, and artificial intelligence. Some of the available research results have already been marshaled in chapter 2. More will be explored in the final three chapters. The conclusion of this chapter—that our familiar self-conception is and always has been a theoretical conception in its own right places all of those results in a deeper perspective.

As we shall see, the network theory of meaning also has major consequences for the vexing epistemological problems explored in the next chapter. We shall turn to those problems after examining one final issue concerning meaning: the *intentionality* of many of our mental states.

Suggested Readings

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* (New York: Routledge & Keegan Paul, 1963); see especially sections 45–63.

Fodor, Jerry, and Chihara, C., "Operationalism and Ordinary Language: A Critique of Wittgenstein," *American Philosophical Quarterly*, vol. 2, no. 4 (1965).

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

Hempel, Carl, and Oppenheim, Paul, "Studies in the Logic of Explanation," *Philosophy of Science*, vol. 15 (1948), part I. Reprinted in *Aspects of Scientific Explanation*, ed. Carl Hempel (New York: Collier-Macmillan, 1965).

Churchland, Paul, "The Logical Character of Action Explanations," *Philosophical Review*, vol. LXXIX, no. 2 (1970).