Physicalism and the Indeterminacy of Translation

Michael Friedman
Harvard University

I

Quine's thesis of the indeterminacy of translation is probably the most well known and most widely discussed thesis in contemporary philosophy. It seems to me, however, that despite its widespread discussion both the content of the thesis and the arguments for it remain relatively unclear. I think the main reason for this unclarity is that the thesis characteristically takes two different forms: (a) an epistemological form, in which it concerns the relation between translation manuals and the possible evidence or data which we use to choose between such manuals, and (b) an ontological form, in which it concerns the relation between translation manuals and "the totality of facts" or "the whole truth about nature"—it is this latter form which is expressed by saying that there is no fact of the matter about correct translation. Because the indeterminacy thesis takes two different forms, it is difficult to evaluate arguments for it. Is a given argument an argument for indeterminacy in the first sense or the second sense? Do arguments for indeterminacy in the first sense support indeterminacy in the second sense? Etc.

The first form of the indeterminacy thesis treats the issue from an epistemological or methodological point of view. The problem is the relationship between our data and methods for selecting translation manuals, on the one hand, and the translation manual we ultimately select, on the other. In this form, the thesis makes the claim that our data and methods do not determine a unique choice of translation. There will always be incompatible translation manuals equally well supported by the totality of our evidence. This epistemological reading of the indeterminacy thesis is supported by such passages from Quine as the following:
[T]he linguist’s finished jungle-to-English manual has as its net yield an infinite semantic correlation of sentences .... Most of the semantic correlation is supported only by analytical hypotheses, in their extension beyond the zone where independent evidence ... is possible. That these unverifiable translations proceed without mishap must not be taken as pragmatic evidence of good lexicography, for mishap is impossible. ([5]: 71.)

Yet one has only to reflect on the nature of possible data and methods to appreciate the indeterminacy. Sentences translatable outright, translatable by independent evidence of stimulatory occasions, are sparse and must woefully under-determine the analytical hypotheses on which the translation of all further sentences depends. ([5]: 72.)

Many discussions of the indeterminacy thesis have assumed that the main philosophical problem raised by Quine’s treatment of translation is this epistemological or methodological one. Thus, we have all heard philosophers debate about what are the correct “constraints” on the methodology of translation, whether the “principle of charity” is a good methodological principle, etc. Yet this methodological debate, interesting as it is in its own right, does not address itself to the whole problem—or even the main problem—raised by Quine’s work. And this is for two reasons. First, Quine repeatedly states that his point is not an epistemological one. He is not claiming merely that we can never know which is the correct translation, but that there is no correct translation that we can either know or fail to know. There is no fact of the matter as to which translation is correct or incorrect:

The point is not that we cannot be sure whether the analytical hypothesis is right, but that there is not even ... an objective matter to be right or wrong about. ([5]: 73.)

Second, as is well known, Quine holds a more general thesis of the underdetermination of theory. He holds that not only translation theory but all of science is epistemologically undetermined by our evidence. There will always be incompatible total scientific theories equally well supported by the totality of our evidence. But the indeterminacy thesis is supposed to distinguish translation theory from the rest of science; it is supposed to point out a special defect in such linguistic notions as meaning and reference. For example, in an
often referred to passage from his reply to Chomsky in *Words and Objections*, Quine says:

Though linguistics is of course a part of the theory of nature, the indeterminacy of translation is not just inherited as a special case of the under-determination of our theory of nature. It is parallel but additional. Thus, adopt for now my fully realistic attitude toward electrons and muons and curved space-time, thus falling in with the current theory of the world despite knowing that it is in principle methodologically under-determined. Consider, from this realistic point of view, the totality of truths of nature, known and unknown, observable and unobservable, past and future. The point about indeterminacy of translation is that it withstands even all this truth, the whole truth about nature. This is what I mean by saying that, where indeterminacy of translation applies, there is no real question of right choice; there is no fact of the matter even to within the acknowledged under-determination of a theory of nature. ([7]: 303.)

Thus, it is clear that Quine intends to be making something more than an epistemological claim in his thesis of the indeterminacy of translation. He wants to say that not only is translation not determined by all our evidence, it is not even determined by all the facts there are—not determined by all the truths about nature. Now, this strong claim is trivially false if we don’t impose some kind of limitation on what can be part of the “totality of truths of nature”—otherwise, we can simply count our preferred translation manual as part of this totality. In Quine’s case, the necessary limitation derives from his physicalism. He believes that physics is our most basic and fundamental theory of the world (our “ultimate parameter”) and that the totality of truths about the entities dealt with by physics represents the totality of truths that there are. Therefore, since Quine believes that the totality of truths of nature = the totality of truths of physics, the ontological version of the indeterminacy thesis amounts to the claim that translation is not determined by the set of truths of physics. Translation theory is not determined by physical theory in the way chemical theory, for example, is determined by physical theory.

In this paper I concentrate mainly on the ontological version of the indeterminacy thesis. After attempting to further clarify the distinction between the two forms of the thesis in Section II, I go on in Section III to argue that Quine has not
provided us with a reason for thinking that translation theory is undetermined (in the relevant sense) by the totality of physical facts. Quine has not provided us with a reason for thinking that linguistic theory is different from any other higher-level theory—like chemistry or biology—in this respect. In short, I try to show that one can accept Quine’s physicalism without accepting the indeterminacy thesis. If I am right, the issue of indeterminacy is therefore independent of the dispute between Quine and the (anti-physicalistic) Frege-Church-Carnap tradition in semantics. As I suggest in Section IV, however, there is certainly an important historical connection between Quine’s rejection of the Frege-Church-Carnap tradition, on the one hand, and his advocacy of the indeterminacy thesis, on the other.

II

As stated above, neither form of the indeterminacy thesis is very clear. What is the “totality of our evidence” and the “totality of truths of physics”? What does it mean for either of these totalities to “determine” a theory? Following some hints of Quine, I will try to state the two theses more precisely.

For simplicity, I adopt the fiction that our language is an interpreted first-order language whose domain of interpretation is the set of space-time points—so all predicates are predicates of space-time points. I also suppose that our physical theory is a definite first-order theory, with a definite set of primitive predicates. Finally, I will assume that our language contains a distinguished set of observational predicates, picked out, for example, as in Quine [6]: 86-89.

Now specify, for each observational predicate ‘Ox’ and each space-time point q, whether ‘Ox’ is true of q. (I ignore many-place predicates, but they can obviously be treated analogously.) Such a specification will be said to be a specification of all possible evidence. What is it for a theory to be epistemically determined by such a specification? Sometimes Quine writes as if ‘epistemically determines’ just means ‘entails’—e.g., when he states the doctrine of the underdetermination of theory in the form: there are always incompatible theories which are each compatible with all possible evidence (cf., e.g., [8]: 179). In this form, the doctrine is of course trivially true for any theory that essentially contains non-
observational predicates. At other times, Quine states the doctrine in the stronger form: there are incompatible theories which are each compatible with all possible evidence and equally in accordance with "the ideal organon of scientific method" (cf., e.g., [5]: 22). This stronger form of the doctrine suggests that Quine uses 'epistemically determine' in such a way that a theory is epistemically determined just in case it is compatible with all possible evidence and there is no incompatible theory which is equally in accordance with "the ideal organon of scientific method." This latter sense of 'epistemically determine' is clearly more interesting, although vaguer, than the first; but for present purposes it doesn't really matter which sense we use. So from now on I'll let 'epistemically determine' be ambiguous as between the stronger and the weaker sense.

What about determination by the truths of physics? (Call this ontological determination.) We can start by specifying for each primitive predicate of physics 'Px' and each space-time point q, whether 'Px' is true of q (again, I ignore many-place predicates). Such a specification will be said to be a specification of the totality of truths of physics. What is it for a theory to be (ontologically) determined by such a specification? First of all, it is clear that one way a theory can be determined by physics is by being reducible to physics in the classical sense. A theory is reducible to physics in this sense if each predicate of the theory is coextensive with a predicate of physics and the laws of physics constrain the corresponding physical predicates to satisfy the theory. To be a little more precise: Let T1 be physics, T2 the theory to be reduced, and 'F1x', 'F2x', ..., 'Fnx' the primitive predicates of T2. (I continue to ignore many-place predicates.) Let a physical interpretation be a mapping α which associates each 'Fi' with an open sentence containing only physical predicates, α('Fi') = 'Ai'. For any sentence containing only predicates from among 'F1x', 'F2x', ..., 'Fnx', we can define truth under the interpretation α and satisfaction under α—they are defined just like satisfaction and truth, except that the clause for atomic formulas now reads:

A sequence σ satisfies 'Fi' under α iff σ satisfies α('Fi').

Thus, T2 is strongly (classically) reducible to T1 if there is a physical interpretation α of T2's primitive predicates such that
for each predicate ‘$F_i x$’ and each space-time point $q$, ‘$F_i x$’ is true of $q$ just in case $\alpha(\{F_i x\})$ is true of $q$ (coextensiveness) and $T_2$ comes out true under $\alpha$ in every model of $T_1$.

However, it is also clear that there are more general ways in which a theory can be determined by physics. One such way emerges from the debate between identity theorists and functionalists in the philosophy of mind. The identity theorist argues that the relationship between mental and physical predicates is to be construed analogously to “theoretical identifications” like water = $\text{H}_2 \text{O}$ and temperature = mean kinetic energy. He thinks that mental predicates are explicitly definable in terms of physical predicates and, therefore, that psychology is strongly reducible to physics. The functionalist, on the other hand, argues against explicit definability. He argues, for example, that there is no one physical state that is always present when someone is in pain. To say that someone is in pain is to say that he is in some state or other which has the functional role characteristic of pain, and there are an indefinite number of distinct physical states that can play that functional role. Nevertheless, the functionalist thinks that each particular instance of a mental state is identical with some particular physical state; mental states are “realized” by particular physical states.

Can we make the kind of physical determination which the functionalist argues for more precise? The crucial difference between the identity theorist and the functionalist is that the identity theorist thinks that each mental predicate corresponds to a single physical predicate, while the functionalist thinks that each mental predicate corresponds to an indefinite number of physical predicates. Thus, it is natural to generalize the classical notion of reduction by associating each non-physical predicate, not with a single physical predicate, but with a set of physical predicates: Let ‘$F_1 x$', ‘$F_2 x$', ..., ‘$F_n x$' again be the primitive predicates of the theory to be reduced. Let a physical realization be a mapping $\beta$ which associates each ‘$F_i x$’ with a set of open sentences containing only physical predicates, $\beta(\{F_i x\}) = \{A_1^i x'$, $A_2^i x'$, ...$\}$. For any sentence containing only predicates from among ‘$F_1 x$', ‘$F_2 x$', ..., ‘$F_n x$', we can define truth under the realization $\beta$ and satisfaction under $\beta$—they are defined just like satisfaction and truth, except that the clause for atomic formulas now reads:
A sequence $\sigma$ satisfies \( F_i x \) under $\beta$ iff there exists an $A_i x' \in \beta('F_i x')$ such that $\sigma$ satisfies $A_i x'$.

Let us now define weak reducibility analogously to the above definition of strong reducibility: A theory is *weakly reducible* to physics if there is a physical realization $\beta$ of its primitive predicates such that for each predicate $F_i x'$ and each space-time point $q$, $F_i x'$ is true of $q$ just in case some $A_i x'$ in $\beta('F_i x')$ is true of $q$ ("$F_i x'$ is not coextensive with any single physical predicate, but rather with a "disjunction"—possibly infinite—of physical predicates) and in every model of physics the theory comes out true under $\beta$.

I don’t know if there are other kinds of physical determination besides these two kinds, if there are other kinds of physical determination besides what I have called weak and strong reduction. However, I hope that it will become clear in the course of my discussion that it is these two kinds of physical determination which are most relevant to what Quine says about the indeterminacy of translation. So from now on I will interpret ‘physical determination’ as ‘weak or strong reduction’ and interpret physicalism as the doctrine that all "respectable" predicates and theories must be weakly or strongly reducible to physics. Quine’s complaint against linguistics and translation theory, then, is that they unlike other higher-level theories—fail to be (strongly or weakly) reducible to physics.

If all this is a fair interpretation of what Quine means by “being determined by all possible evidence” and “being determined by the totality of truths of nature” respectively, it follows that the two forms of the indeterminacy thesis are very different. First of all, the two forms of the thesis concern relations between different terms. The epistemological form of the thesis concerns a relation between translation theory and a specification for each *observational* predicate ‘$O_x$’, and each space-time point $q$, whether ‘$O_x$’ is true of $q$. The ontological form of the thesis concerns a relation between translation theory and an analogous specification for each *physical* predicate. Secondly, the relations referred to are different. The first version concerns a relation of epistemic determination (understood ambiguously as above); the second version concerns a non-epistemic relation of (weak or strong) reducibility.
Furthermore, although if we interpret ‘epistemically determines’ as ‘entails’ and assume that all observational predicates are physically determined, the ontological version of the thesis implies the epistemological version; the epistemological version definitely does not imply the ontological version. From the fact that a given theory is not epistemically determined by all observational truths, it by no means follows that it is not ontologically determined by all physical truths. Indeed, if Quine is right about the underdetermination of theory, there exist theories today which are ontologically determined but not epistemically determined. Chemical theory, for example, is, like all theory, not uniquely determined by all possible evidence; yet it is (strongly?) reducible to physics and is therefore ontologically determined—there is a fact of the matter about chemistry. Consequently, it would seem that arguments for the weaker, epistemological, version of the thesis do not go very far, if at all, towards supporting the stronger, ontological, version. Readers of *Word and Object* are therefore justifiably puzzled, I think, by the fact that most of Quine’s discussion of translation is occupied with epistemological and methodological issues—issues which have no obvious connection with the more interesting question of whether there is a fact of the matter about translation, whether translation theory is physically determined. In fact, it is hard to find any passage in Quine’s writings which is clearly an argument for the ontological version of the indeterminacy thesis. Nevertheless, I think Quine does have an argument for the claim that translation theory is not physically determined. I will examine it in the next section.

### III

Quine’s clearest and most explicit argument for the indeterminacy thesis is found in “On the Reasons for Indeterminacy of Translation” [8]. There he uses the doctrine of underdetermination of theory to argue for indeterminacy in the following way. If underdetermination is true, there can be two incompatible total theories $A$ and $B$ formulable in our language which are both equally well supported by all possible true observation sentences of our language. Therefore, if we translate a foreign speaker’s observation sentences—by matching their stimulus meanings with the stimulus meanings of our observation sentences—and thus determine which observation sentences (in
our language) he accepts, we still haven’t determined which theoretical sentences (in our language) he accepts. We are free to translate his theoretical sentences in such a way that he comes out as holding either theory A or theory B, because A and B are supported equally well by all possible observation sentences in the foreign speaker’s language:

Insofar as the truth of a physical theory is underdetermined by observables, the translation of the foreigner’s physical theory is underdetermined by translation of his observation sentences. If our physical theory can vary though all possible observations be fixed, then our translation of his physical theory can vary though our translations of all possible observation reports on his part be fixed. Our translation of his observation sentences no more fixes our translation of his physical theory than our own possible observations fix our own physical theory. ([8]: 179-80.)

This argument is a little more complicated than it first appears, because of the possible ambiguity of ‘fix’. On the one hand, the observation sentences we accept do not epistemically determine for us (fix) the theoretical sentences we accept. On the other hand, the observation sentences the foreigner accepts do not epistemically determine for him (fix) the theoretical sentences he accepts. The conclusion Quine wants to derive from this is that our translation of the foreigner’s observation sentences does not epistemically determine for us (fix) our translation of his theoretical sentences. It would seem that some such principle as: if \( S_1 \) does not epistemically determine \( S_2 \) for person \( P_1 \), then ‘\( P_1 \) believes \( S_1 \)’ does not epistemically determine ‘\( P_1 \) believes \( S_2 \)’ for person \( P_2 \), is required. In any case, I will assume that some such principle is plausible and that consequently Quine’s argument that our translation of a foreign speaker’s observation sentences does not epistemically determine (for us) our translation of his theoretical sentences is valid (given the assumption of the underdetermination of theory, of course).

How does the above argument support the indeterminacy thesis? Consider first the epistemological form of the thesis. The conclusion of the above argument is that our translation of a foreigner’s observation sentences does not epistemically determine our translation of his theoretical sentences. What the (epistemological) indeterminacy thesis says is that all possible evidence does not epistemically determine our translation of his
theoretical sentences. Now there are two important things to notice about the relation between all possible evidence, on the one hand, and the translation of observation sentences, on the other. First, the translation of observation sentences itself goes beyond all possible evidence in the strict sense—it is not determined by the totality of observational facts. Let us call a predicate ‘$Bx$’ a behavioral predicate if ‘$Bx$’ is true of $q$ just in case there is an organism $O$ at $q$, a stimulation $\sigma$, and an item of observable behavior $\beta$, such that if $O$ were to receive $\sigma$, $O$ would emit $\beta$. That is, ‘$Bx$’ is a behavioral predicate just in case it attributes a behavioral disposition to some organism. In *Word and Object*, Quine shows that the translation of observation sentences is determined by behavioral facts, by the distribution of true behavioral predicates over space-time points. However, behavioral predicates are not themselves observational, nor are they determined by all possible observational facts (cf., e.g., [5]: 222-26). The attribution of a behavioral predicate essentially involves a (low-level) theoretical inference. Thus, Quine’s conclusion is strictly stronger than the epistemological indeterminacy thesis as I have stated it. Not only is translation underdetermined by the observational facts, it is not even determined by the totality of observational and behavioral facts.

Secondly, however, it is not clear that the translation of observation sentences exhausts the available evidence. Quine has shown (let us assume) that the translation of observation sentences does not determine the translation of theoretical sentences, and that the translation of observation sentences is determined by the observational + behavioral facts. To get the required conclusion—that the translation of theoretical sentences is not determined by the observational + behavioral facts—we need an additional premise to the effect that the translation of observation sentences is all the evidence that could possibly be relevant to the translation of theoretical sentences. And this is at least not obviously true—e.g., there could conceivably be non-linguistic behavioral facts that would be relevant to the translation of theoretical sentences. However, Quine himself certainly thinks that it is true:

In order . . . to construe the foreigner’s theoretical sentences we have to project analytical hypotheses, whose ultimate justification is substantially just that the implied observation sentences match up. ([8]: 179.)
Be this as it may, I will again assume that Quine has at least made this crucial claim plausible and that the above argument does support the epistemic version of the indeterminacy thesis (strengthened to include behavioral as well as strictly observational evidence as above).

What then of the non-epistemological form of the indeterminacy thesis? The first difficulty is that the argument I have been considering appears to deal exclusively with epistemic determination. However, we can immediately connect it up with ontological (reductive) determination if we assume that if a set of truths $\mathcal{F}$ ontologically determines a theory $T$, then $\mathcal{F}$ could, in principle anyway, epistemically determine $T$ (since the truth and falsity of each sentence in $T$ is settled by $\mathcal{F}$); and that consequently if $\mathcal{F}$ could not in principle epistemically determine $T$, then $\mathcal{F}$ does not ontologically determine $T$. The argument then gives us the result that the totality of observational + behavioral facts does not ontologically determine the translation of theoretical sentences (where determination by observational + behavioral facts is defined just like determination by physical facts, except that ‘observational or behavioral predicate’ is substituted everywhere for ‘physical predicate’). The translation of theoretical sentences is not (strongly or weakly) reducible to facts about behavior, to facts about dispositions to assent and dissent given various sensory stimulations.

It seems to me that this last conclusion is both extremely plausible in its own right and quite well supported by Quine's arguments. However, it is still very far from the strong, ontological, form of the indeterminacy thesis. For, according to that form of the thesis, translation (of theoretical sentences) is not ontologically determined by any physical facts at all. And there are a lot more physical facts than facts about dispositions to assent and dissent given various sensory stimulations; there are a lot more physical facts than behavioral facts. (The set of observational + behavioral predicates is only a small fraction of the set of physically determined predicates.) Therefore, it would seem that if the argument we have been considering is to have any relevance to the strong indeterminacy thesis, we have to adopt some form of behaviorism about linguistics and translation theory. We have to suppose that the only physical facts that could be relevant to the reduction of translation theory to physics are behavioral facts. Furthermore, it seems
clear that Quine himself holds such a behaviorist view of linguistics:

When with Dewey we turn thus toward a naturalistic view of language and a behavioral view of meaning, what we give up is not just the museum figure of speech. We give up an assurance of determinacy. . . . When . . . we recognize with Dewey that "meaning . . . is primarily a property of behavior," we recognize that there are no meanings, nor likenesses nor distinctions of meaning, beyond what are implicit in people's dispositions to overt behavior. For naturalism the question whether two expressions are alike or unlike in meaning has no determinate answer, known or unknown, except insofar as the answer is settled in principle by people's speech dispositions, known or unknown. ([6]: 28-29.)

If this is correct, then attempts like Harman's [2] to separate the indeterminacy thesis from behaviorism are misguided. Not only does Quine explicitly link his thesis to behaviorism, but without some sort of behavioristic assumption to the effect that the only facts that could be relevant to translation are behavioral facts, Quine's argument for the strong indeterminacy thesis is simply a non sequitur—since the totality of physical facts ≠ the totality of behavioral facts. However, Harman is certainly right to insist that Quine's argument is not based on a general assumption of philosophical behaviorism—an assumption that all mental facts must be behavioristically reducible. Rather, Quine has special reasons for thinking that behaviorism is true for specifically linguistic facts. I will examine these reasons in what follows.

Let me first briefly review the situation. Quine wants to claim not only that there can be no grounds for choosing a unique correct translation, but that there is no objective fact to be right or wrong about. And what he means by this latter claim seems to be that there are no physical facts that determine what the correct translation is; translation theory is not determined by physics. However, the argument we have considered so far at most gives us the conclusion that there are no behavioral facts that determine what the correct translation is; translation theory is not determined by purely behavioral facts. Therefore, it is natural for someone (like myself) who agrees with Quine's physicalism but not with the indeterminacy thesis to suppose that there are other, physical but non-behavioral, facts which do determine translation, that there are non-behavioral physical facts to which translation theory is reducible. Plausible candi-
dates for such facts are facts relating our uses of words to our internal (physiological) states, facts relating our uses of words to external physical objects (as so-called "causal" theories of reference suggest), etc. If Quine is to make a convincing case for indeterminacy, he must give us some reason to think that such non-behavioral facts are not relevant to reducing translation theory and that, consequently, since behavioral facts do not suffice to determine translation, nothing does.

I think Quine has several related arguments for the claim that only behavioral factors are relevant. The first appeals to the fact that one learns language on the basis of the observable behavior of others: In learning a language, all the facts one has access to are facts involving the observable behavior of speakers of the language. One doesn't have access to non-behavioral facts, such as facts about the speakers' internal neural states. If there were linguistic facts which were determined by non-behavioral factors (like internal neural states) but not by behavioral factors, these facts could not be learned in the course of acquiring a language. Therefore, all linguistic facts must be determined solely on the basis of observable behavior. This kind of argument is suggested by passages like the following:

Meanings are, first and foremost, meanings of language. Language is a social art which we all acquire on the evidence solely of other people's overt behavior under publicly recognizable circumstances. Meanings, therefore, those very models of mental entities, end up as grist for the behaviorist's mill. Dewey was explicit on the point: "Meaning . . . is not a psychic existence; it is primarily a property of behavior." ([6]: 26-27.)

Language is a social art. In acquiring it we have to depend entirely on intersubjectively available cues as to what to say and when. Hence there is no justification for collating linguistic meanings, unless in terms of men's dispositions to respond overtly to socially observable stimulations. An effect of recognizing this limitation is that the enterprise of translation is found to be involved in a certain systematic indeterminacy . . . . ([5]: ix.)

At first sight, this argument appears to rest on a simple confusion between epistemic determination and ontological determination. From the premise that linguistic facts are epistemically determined by behavioral facts alone, it by no means follows that they must be ontologically determined by behavioral facts alone (unless we adopt an extremely strong version of 'epistemically determines'). However, we can easily
rephrase the argument in a way which doesn’t confuse the two forms of determination (which doesn’t require us to assume that epistemically determines = entails) as follows: All linguistic facts must be epistemically determined by behavioral factors alone, since these are all that are available in acquiring a language. Anything not epistemically determined by behavioral factors cannot be a linguistic fact, cannot be a fact about meaning. But we know from the epistemic indeterminacy thesis that the translation of theoretical sentences is not epistemically determined by behavioral factors; behavioral factors do not epistemically determine a unique, distinguished meaning for each theoretical sentence. Therefore, there are no linguistic facts determining the translation of theoretical sentences. Non-behavioral factors, like internal neural states, are consequently irrelevant to the translation of theoretical sentences. Non-behavioral facts cannot be linguistic facts.

I think what is wrong with this argument is the premise that linguistic facts must be epistemically determined by behavioral facts. It is certainly true that the only evidence we have available to us in acquiring language is the observable behavior of others; we learn language by something like an inference from other speakers’ behavior. But it doesn’t follow that this inference has to be epistemically determined (especially if we interpret ‘epistemically determines’ in a very strong sense). On the contrary, it is plausible to view this inference as analogous to an inductive or theoretical inference and, as such, subject to the usual inductive or theoretical underdetermination. (In this connection, remember that even the attribution of behavioral dispositions involves a theoretical inference.) We learn language on the basis of other people’s observable behavior, and we learn facts about electrons on the basis of the behavior of ordinary observable objects. In neither case is the former epistemically determined (much less ontologically determined) by the latter. Thus, the fact that we learn language on the basis of observable behavior gives us no reason to think that all linguistic facts must be epistemically (or ontologically) determined by facts about behavior.

A second Quinean argument again starts from the fact that one learns language by a social process which establishes conditioned relations between one’s linguistic behavior and sensory stimulations. According to this argument, one has successfully mastered a language—one can communicate in that
language—when the conditioned relations between one’s linguistic behavior and sensory stimulations match up sufficiently with those of other speakers of the language. Now, non-behavioral factors (internal physiological states, “causal” relations to external objects, etc.) can vary arbitrarily as long as these conditioned relations are preserved. Such other factors, therefore, cannot be linguistically relevant; they cannot be relevant to the physical determination of translation. Quine appears to give a version of this argument in the following passage:

The sort of meaning that is basic to translation, and to the learning of one’s own language, is necessarily empirical [=behavioral] meaning and nothing more. A child learns his first words and sentences by hearing and using them in the presence of appropriate stimuli. These must be external stimuli, for they must act both on the child and on the speaker from whom he is learning. Language is socially inculcated and controlled; the inculcation and control turn strictly on the keying of sentences to shared stimulation. Internal factors may vary ad libitum without prejudice to communication as long as the keying of language to external stimuli is undisturbed. Surely one has no choice but to be an empiricist [=behaviorist] so far as one’s theory of linguistic meaning is concerned. ([6]: 81.)

In this passage, Quine refers only to non-behavioral internal facts, but he would presumably make the same point about non-behavioral external facts like “causal” relations with external objects.

I don’t think this argument is convincing. First of all, it at most shows that translation is not strongly determined by non-behavioral physical facts, that linguistic notions like meaning and reference are not explicitly definable in terms of non-behavioral physical predicates. The possibility remains open that such linguistic notions are weakly determined by non-behavioral physical predicates, and that translation theory is therefore weakly reducible to physics. Suppose it were true that non-behavioral factors vary freely from speaker to speaker, and even vary from occasion to occasion with the same speaker. It would follow that linguistic notions could not be defined in terms of such non-behavioral factors, for different non-behavioral factors would underlie the same linguistic behavior at different times. However, we could still consistently suppose that each instance of linguistic behavior is determined by some physical, but non-behavioral, fact about the speaker. We could suppose that linguistic behavior is weakly, but not strongly,
determined by physical facts. To return to the analogy with the philosophy of mind, Quine's argument shows at most that notions like meaning and reference have to be understood functionally relative to physical predicates, not that they have to be understood behavioristically. Note that the functionalist is no more a behaviorist than the identity theorist is. He does not expect that any interesting part of human behavior (including linguistic behavior) can be understood in purely behavioral terms.

Secondly, I don't think that Quine's argument even establishes that translation is not strongly determined by non-behavioral factors, that linguistic notions are not explicitly definable in terms of physical ones. For, in this connection, the crucial question is not whether non-behavioral factors could in principle vary arbitrarily, but whether they do in fact. Imagine for a moment that translation theory is at least weakly reducible to physics, so that there are physical predicates which realize such linguistic notions as meaning and reference. The only remaining question for the issue of strong reducibility is whether the physical predicates which realize the linguistic predicates vary from case to case or not. If the physical predicates vary sufficiently, linguistics will fail to be strongly reducible to physics; if not, linguistics will be strongly reducible. Nothing at all follows from the mere possibility that the physical predicates vary arbitrarily.

Imagine someone arguing that chemical structure cannot be relevant to explaining the observable properties of tap water, river water, sea water, etc., because it is in principle possible for such chemical structure to vary arbitrarily from case to case as long as the same observable properties are preserved. This would obviously be a non sequitur. For, first, even if chemical structure does vary arbitrarily, it may still be relevant to explaining the observable properties of water. It may weakly determine the properties of water even if it doesn't strongly determine them. Second, the issue is not what is in principle possible but what is in fact the case. And, as a matter of fact, all instances of water do share a similar chemical structure, a chemical structure which explains the similarity in their observable properties. I think the case of linguistics and translation theory is analogous. The issue is whether such non-behavioral factors as internal physiological states and external relations with physical objects are relevant to explain-
ing linguistic behavior, and thus relevant to reducing translation theory to a physical basis. From the claim that such factors may vary arbitrarily while linguistic behavior is preserved, it by no means follows that they do. And even if non-behavioral factors do vary arbitrarily, it doesn’t follow that they can’t be relevant to explaining linguistic behavior and physically determining linguistic theory.

Thus, it seems to me that the central issue underlying the debate over whether there is a fact of the matter about translation is an empirical one—it is not something that can be settled by philosophical argument. The issue is whether there are non-behavioral physical facts which are sufficient to (weakly or strongly) reduce linguistics and translation theory to physical science. At best, therefore, what Quine is doing is betting on the future course of science. Quine is betting that science will not uncover such facts and that consequently only behavioral facts will prove explanatorily relevant. Quine’s opponents are betting that science will uncover such facts. Given Quine’s general philosophical orientation, it should perhaps not be too disturbing that the thesis of indeterminacy of translation turns out to be a guess as to the future progress of science. What is more disturbing is that it seems we have been given no reason for preferring Quine’s guess to its opposite.

IV

Quine’s attitude towards most of linguistic theory, particularly towards such semantical notions as meaning and reference, is basically one of skepticism. This skepticism has both epistemological and ontological motivations. On the one hand, Quine despairs of finding evidence that would enable one to decide between alternative, mutually incompatible imputations of meanings or of referents. On the other hand, he despairs of finding physical facts to which such notions as meaning and reference can be (weakly or strongly) reduced—in the way that, for example, there are physical facts to which such chemical notions as valence can be reduced. And, for a physicalist, if there are no physical facts underlying a given area of discourse, there are no objective facts at all. As I indicated above, I think that the second, ontologically motivated, form of skepticism is the most serious form, especially since I am in sympathy with Quine’s physicalism. I also think—and I have tried to show
above—that Quine’s direct arguments for this kind of skepticism are far from conclusive. Nevertheless, it seems to me that Quinean skepticism becomes both understandable and plausible when viewed in its historical context.

In pre-Quinean philosophy of language, we find two main traditions. First, there is the verificationist tradition, which views the connection between language and sense experience as the most important semantical relation. Second, there is the Frege-Church-Carnap tradition, which explains the semantical properties of language by postulating such entities as concepts and propositions—where these latter are thought of as either irreducibly mental entities or abstract entities. Now, as a naturalist, Quine clearly could not be happy with this latter type of approach, because of its promiscuous postulation of scientifically unrespectable entities. (Quine, of course, is not opposed to postulating abstract entities per se. He thinks—reluctantly—that we should postulate all mathematical entities necessary for the mathematics of physics. Because of his physicalism, however, he would oppose the postulation of any special abstract entities—like concepts or propositions—which are not part of the ontology required by physical science.) On the other hand, the prospects for making scientific sense out of the verificationist approach to language appeared distinctly better. In fact, I think that one of Quine’s major contributions is the way he shows how to put the verificationist tradition into a naturalistic setting—by means of the behavioral concepts of stimulus meaning, observation sentence, etc. However, Quine also saw what none of the earlier verificationists did, namely, that a consistent and physically respectable verificationism (i.e., behaviorism) leads to skepticism about semantics, leads to the indeterminacy of translation and the inscrutability of reference (cf. [6]: 80-81).

If this is correct, Quine’s naturalism led him to verificationism, and his naturalized verificationism led him to skepticism. Since I think that a naturalistic approach to language is salutary, I think that if we want to reject Quinean skepticism, we should reject his verificationism, not his physicalism. According to Quine, the only semantically relevant physical facts are behavioral ones—facts relating linguistic behavior to sensory stimulation—and these facts are insufficient to make sense of such traditional linguistic notions as meaning and reference. If we want to reject skepticism but retain naturalism,
we have to look for other, possibly non-behavioral, physical facts to which these semantic notions can plausibly be (weakly or strongly) reduced. This latter type of program can be looked at as a type of "reductive physicalism" in contrast to Quine's "eliminative physicalism." (As these labels suggest, I view this issue as importantly analogous to the dispute between "reductive materialism" and "eliminative materialism" in the philosophy of mind—and I classify both identity theorists and functionalists as "reductive materialists", since both think that mental states are realized by physical states. The "reductive materialist" thinks that mental states and properties are identified with, or at least realized by, physical states and properties; the "eliminative materialist" thinks that mental states and properties do not have such a physical basis and that they should therefore be eliminated from scientific discourse. Similarly, the "reductive physicalist" in semantics thinks that semantic notions are scientifically respectable, for they are (weakly or strongly) reducible to purely physical ones; the "eliminative physicalist"—Quinean skeptic—thinks that semantic notions are not physicalistically reducible and that they are therefore not scientifically respectable.)

What would such a "reductive physicalism" look like? I think that one kind of answer is suggested by recent work on so-called "causal" theories of reference. According to this type of view, there are physical relations between our uses of words and physical objects in virtue of which words refer to the objects that they do refer to. There are physical relations between our use of the word 'Socrates' and Socrates in virtue of which 'Socrates' refers to Socrates; there are physical relations between our use of the word 'red' and red things in virtue of which 'red' applies to red things, etc. This type of view contrasts with the traditional Frege-Church-Carnap theory according to which reference is determined by the (non-physical) sense of an expression. It also contrasts with the Quinean skeptical approach according to which the only semantically relevant physical relations between words and non-linguistic entities relate our uses of words to sensory stimulations, stimulus meanings. Since different referents can yield the same stimulus meanings, we end up with the doctrine of inscrutability of reference. On the other hand, if we look at things from the point of view of a broader physicalism, the possibility of referential determinacy is restored. Thus, for
example, if we suppose that the reference relation can be reduced to a certain physical relation $R$, we may find that $R$ holds between a given tribe’s use of the word ‘gavagai’ and rabbits, while $R$ does not hold between their use of the word ‘gavagai’ and rabbit stages. (An analogous possibility exists if reference is realized, not by a single physical relation $R$, but by a set of physical relations $\{R_1, R_2, \ldots\}$.) This would give an objective sense to saying that ‘gavagai’ refers to rabbits but not rabbit stages.

It is tempting for critics of this kind of physicalistic view of reference to adopt a “put up or shut up” attitude, to demand more information about the nature of this mysterious relation $R$ (or about the set of mysterious relations $\{R_1, R_2, \ldots\}$ if weak reducibility is preferred) before taking the view seriously. This attitude, I think, is mistaken. Spelling out the precise nature of such a relation (or set of relations) is not a job for philosophers. Actually carrying out a physicalistic reduction of semantic notions like reference, if it is to be carried out at all, is a task for empirical scientists—for linguists and psychologists. (For this reason I don’t think it is particularly helpful to call the view I am considering a “causal” theory of reference. Not only is the notion of causal connection itself notoriously unclear, but we are simply not in a position today to have much of a theory about the reference relation—just as we are not in a position today to have much of a theory about the brain processes underlying our psychological states. What is important—and I think the contrast with Quine’s views shows this—is that reference is seen as a physical relation between our uses of words and the objects they refer to.) What a philosophical proponent of this kind of program can do is: (1) try to disarm philosophical arguments purporting to show that the intended reduction is in principle impossible (like the Quinean argument I considered above purporting to show that only behavioral facts could be semantically relevant), (2) argue that it is plausible that semantic notions should be so reducible, (3) argue that looking at semantic notions from a “reductive physicalist” point of view illuminates traditional philosophical issues—notably the concept of truth (cf. [1]), (4) try to show that such a point of view contrasts favorably with other philosophical attitudes towards semantics (cf., e.g., [3], [4]). Again, I think the analogy with the philosophy of mind is helpful. It is not up to the philosophical defender of “reductive materialism” to say
precisely which neurophysiological states are to be identified with (are to realize) which mental states. What he can do, rather, is argue for the possibility and plausibility of the envisioned reduction, argue that looking at things from his point of view illuminates traditional disputes in the philosophy of mind, argue that his point of view contrasts favorably with other philosophical views such as dualism and analytical behaviorism, etc.

In conclusion, I think that Quine’s skeptical attitude towards semantics is best seen as falling naturally into one of three important traditions in the philosophy of language. The verificationist tradition, of which Quine’s views are a natural extension, sees the connection between language and sense-experience as centrally, and perhaps exclusively, important for semantics. The Frege-Church-Carnap tradition sees the postulation of special abstract entities such as concepts and propositions as necessary for understanding language. The theory-of-reference type of approach sees the (physical) relation between language and physical objects as centrally important. My own sympathies are obvious. Because I share Quine’s naturalism, I find the Frege-Church-Carnap tradition barren and unfruitful, while the theory-of-reference approach impresses me as the most promising. Nevertheless, it is certainly true that I have here given the reader no grounds for choosing one of the three approaches over the others. What I have attempted to do, however, is show that Quine has not given us a reason to think we have no choice but skepticism.¹

REFERENCES

NOTE

1 I am indebted to discussions with and advice from Hartry Field, Warren D. Goldfarb, Robert Harding, and David Hills.